EVALUATION OF POTASSIUM STRESS RESPONSES AND IDENTIFICATION OF NOVEL RT-qPCR REFERENCE GENES IN THE HALOARCHAEON, *HALOARCULA MARISMORTUI*

by

Matthew W. Jensen

B.Sc., University of Northern British Columbia, 2009

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
MATHEMATICAL, COMPUTER AND PHYSICAL SCIENCES
(CHEMISTRY)

UNIVERSITY OF NORTHERN BRITISH COLUMBIA

August 2012

Your file Votre référence
ISBN: 978-0-494-94109-6

Our file Notre référence
ISBN: 978-0-494-94109-6

ABSTRACT

Growth characteristics and stress responses in the halophilic archaeon, *Haloarcula marismortui*, have been poorly investigated and knowledge of the effects of extracellular potassium concentration on halophilic growth is limited. We report the evaluation of cellular generation times across a range of extracellular potassium concentrations to assess the organisms responses to extreme potassium stress. Our results show *Haloarcula marismortui* exhibits an optimal generation time of 4.19 ± 0.14 hours at an extracellular KCl concentration of 100mM. This corresponds to an intracellular K^+ concentration of 2.02M as determined through the use of Induction-Coupled Plasma Mass Spectrometry. Additionally, the validation of several candidate reference genes for use with RT-qPCR studies is reported. Five reference genes (*16S rRNA*, *rpoB*, *pykA*, *polA*, and *rpoA*) have been confirmed as being stably expressed in accordance with the Minimal Information for the Publication of Quantitative PCR Experiments (MIQE Guidelines) across several unique halophilic growth conditions.

TABLE OF CONTENTS **Pg**

List of Tables **Pg**

List of Figures **Pg**

List of Symbols and Abbreviations

23% S.W. MGM	23% Salt Water Modified Growth Media
ATP	Adenosine triphosphate
BHB	Bulge-Helix-Bulge motif, found at archaeal intron boundaries
BLAST	Basic Local Alignment Search Tool
bp	Base pairs, referring to nucleic acid bases
CBF5	Centromere-Binding Factor 5
cbiG	A gene encoding Cabalamin biosynthesis protein
cbiH	Gene encoding Precorrin-3B C17-methyltransferase
cbiK	Gene encoding Cobalt chelase thioredoxin
cbiJ	Gene encoding Cobalt-precorrin-6Y C5-methyltransferase
cbiT	Gene encoding Precorrin-8W decarboxylase
cDNA	Complementary Deoxyribonucleic Acid
cobN	A gene encoding Cobalamin biosynthesis protein
C_T	Threshold Cycle
C_q	Quantitative Cycle
DLR	Dynamic Linear Range
DNA	Deoxyribonucleic Acid
DNase	Deoxyribonuclease
dNTP	Deoxyribonucleotide triphosphate
dpa	Gene encoding a signal recognition particle receptor
EDTA	Ethylenediaminetetraacetic Acid
etfB1	Electron transfer flavoprotein beta sub-unit
G+C	Guanine + Cytosine, refers to nucleic acid base pairs
gapB	Gene encoding the GAPDH protein
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
Har. marismortui	Halophilic archaeon native to the Dead Sea
K^+	Potassium Ion
Kb	Kilo-base pairs; 1,000bp
KCl	Potassium Chloride; chemical formula
mA	milliampere
Mb	Mega-base pairs; 1,000,000bp

MIQE	Minimal Information for publication of Quantitative PCR Experiments
mL	Milli-litre, unit of measurement; volume
mM	Milli-molar, unit of measurement; concentration
M-MuLV RT	Moloney Murine Leukemia Virus Reverse Transcriptase
MOPS	3-(N-Morpholino)-propanesulfonic acid
mRNA	Messenger Ribonucleic Acid
M-value	Statistical measure of reference gene stability
NaCl	Sodium Chloride, chemical formula
NCBI	National Center for Biotechnology Information
ng	Nanogram, unit of measurement; mass
No-RT	No Reverse Transcription
NRC-1	Sub-species of *Halobacterium salinarum*
NTC	No-Template Control
OD_{600}	Optical Density at a wavelength of 600 nanometres
oligo-dT	Deoxythymidine oligonucleotide
PCR	Polymerase Chain Reaction
polA1	Gene encoding the DNA polymerase II small sub-unit
polA2	Gene encoding the DNA polymerase II large sub-unit
Poly(A)	Poly-adenylated
Poly(T)	Poly-thymidinylated
Pre-mRNA	Pre-messenger RNA, mRNA prior to intron processing
pykA	Gene encoding the pyruvate kinase protein
qPCR	Quantitative Polymerase Chain Reaction (lacking RT step)
RAP-PCR	RNA Arbitrarily Primed Polymerase Chain Reaction
RNA	Ribonucleic Acid
RNase	Ribonuclease
RNeasy	RNA extraction kit; Qiagen corporation
rpm	Rotations per minute
rpoA	Gene encoding the RNA polymerase alpha sub-unit
rpoB	Gene encoding the RNA polymerase beta sub-unit
RQI	Relative Quality Index
RT	Reverse Transcription

RTL	Buffer supplied as part of RNeasy Kit; Qiagen corporation
RT-qPCR	Reverse-Transcription Quantitative Polymerase Chain Reaction
rRNA	Ribosomal RNA
SucC	Gene encoding Succinyl-CoA synthetase beta chain
SucD	Gene encoding Succinyl-CoA synthetase alpha chain
Sun	Gene encoding Cytosine-C5-methylase
T_a	Annealing Temperature
TAE	Tris-Acetate-EDTA
trkD	Gene encoding a membrane bound potassium ion transporter
trkH	Gene encoding a membrane bound potassium ion transporter
Trk System	A group of proteins responsible for low-affinity potassium ion transport
tRNA	Transfer Ribonucleic Acid
tup1	Gene encoding *tup1*-like transcriptional repressor
UCSC	University of California, Santa Cruz
µL	Micro-litre, unit of measurement; volume
uM	Micro-molar, unit of measurement; concentration
UV	Ultraviolet
% (v/v)	Percent composition by volume
zim	Gene encoding CTAG modification methyl's

Dedication

To my mother, Rachel, and father Willie, your love and encouragement has always been unwavering. You have always supported me in every possible way and stood behind me while I strive to achieve my goals. I truly can not express how grateful I am for this. To my brother, Cory, I have always appreciated your company, advice, and random visits that typically end up with me unexpectedly dropping whatever I'm doing at the time. You have always given me tremendous support and I thank you for making sure I rarely have time for a dull moment.

To my father-in-law, Tony, mother-in-law, Diane, and sisters-in-law, Emily and Maggie, I thank you for your love and support over the past few years and thank you for welcoming me into your family as I set out toward achieving a new goal. I most appreciate your support for your daughter and sister. It has made my countless nights in the lab easier knowing she has always had somebody to talk to while we go without seeing each other for days on end.

Finally, to my wife, Amanda, your love and encouragement has been greater than any other. I can not put into words the gratitude I have for the support you have given me. If it weren't for you I may not have continued this far.

I dedicate this to all of you. I would not be where I am today with out all of your support. Thank you.

And Cory, I'm still not going to get a trade…

Acknowledgement

First and foremost, I would like to thank my supervisor, Dr. Andrea Gorrell, whom I am very fortunate to have been able to study under. I have come to greatly appreciate her mentorship and am grateful for her willingness and patience in allowing me to learn from my own mistakes even when learning has occasionally become quite costly. I am also grateful to her for allowing me to make key decisions in the direction and progression of my research. She has taught me much and has been an amazing role model. I would also like to thank my fellow graduate student, Carly Reinheimer, for her assistance with the initial planning and data collection of this project. To two amazing undergraduate students, Scott Matlock and John Gorman, thank you for your assistance as well. Your desire to get involved with my research and eagerness to learn what you could from me has been truly inspirational. Though the generosity of the above four individuals was instrumental to the completion of this project, I also thank the other members of the Gorrell Lab, past and present, for several years of good conversation and fond memories.

Additionally I would like to thank Dr. Sarah Gray for both sitting on my supervisory committee and providing me with the opportunity to learn RT-qPCR, a technique that I am sure will prove to be highly valuable throughout my career. I am also grateful to Jordie Fraser for taking the time to sit down and talk about the intricacies of RT-qPCR whenever I had become unsure of, or overwhelmed with what felt like, at times, an unending project. Learning from your " nearly militant approach", as Sarah has at times so eloquently referred to it, has shown me the importance of even the most intricate of details while working with this technique.

I am also grateful to Dr. Dezene Huber for sitting on my supervisory committee and providing valuable feedback about my project whenever we had the chance to talk. To Alida Hall, an individual whom I have come to regard as an amazing mentor, colleague, and friend, I would like to thank you for the sound advice you have provided me with over the past several years and appreciate your willingness to sit down and talk about anything and everything. To my closest friends: Meaghan, Dave and Jess, Andrew and Stephanie, Barny, Jon and Carmen, (M)Ike, Tom, Charles, Laura, Anna, Erwin, Randi, Stephane, J.D., and Darla:, as well as my future sister-in-law, Jess, and future brother-in-law, Anthony, I thank you all for your friendship and the support you have given me over the past several years. And finally, to all of my family friends in Fort St. John whom I could always count on for support or a bed to crash on: Mike and Orlanda, Dan and Erin, Danny, Tyler and others, I have always looked forward to coming home for weekend visits.

Thank you all and to the numerous others that have come into my life throughout my years at UNBC.

Chapter One

Introduction

"-a new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it."

- Max Karl Ernst Ludwig Planck, 1858 – 1947

"Convincing ourselves was not the problem. Convincing others was. It would be a hard sell. For reasons I could not understand at the time, literally all biologists believed in "the prokaryote." And it was not your typical scientific belief – always open to question. This was dogma, unshakable doctrine!"

- Carl R. Woese, 1928 -

1.1 The Archaea: Historical Perspectives Behind New Life on Earth

1.1.1 The Procaryote Ideology

The primary distinction between cellular organisms has long been the identification of an organism as either a procaryote or a eucaryote. The phylogenetic classification system has reflected this distinction since the publication of Roger Stanier and C.B. van Niel's paper entitled "The Concept of a Bacterium"[1] in 1962. This now iconic article conveyed the confusion among the microbiological disciplines that persisted due to a lack of sufficient taxonomical classification among the cellular organisms. Stanier and van Niel[1] state, "Any good biologist finds it intellectually distressing to devote his life to the study of a group that cannot be readily and satisfactorily defined in the biological terms; and the abiding intellectual scandal of bacteriology has been the absence of a clear concept of a bacterium." The pair go on to credit their predecessor, Edouard Chatton:

"For a long time, biologists have intuitively recognized that the cell structure of bacteria and blue-green algae is different from that of other organisms, and should be characterized as "primitive"; but a satisfactory description of the difference has proved remarkably elusive. The revolutionary advances in our knowledge of cellular organization which have followed the introduction of new techniques during the past 15 years have changed this situation. It is now clear that among organisms there are two different organizational patterns of cells, which Chatton (1937) called, with singular prescience, the eucaryotic and procaryotic type. The distinctive property of bacteria and blue-green algae is the procaryotic nature of their cells. It is on this basis that they can be clearly segregated from all other protists (namely, other algae, protozoa, and fungi), which have eucaryotic cells."[1]

A more recent review by Jan Sapp[2] points out the work by Edouard Chatton in question, *Titres et Travaux Scientifiques*[3], used the terms "procaryote" and "eucaryote" only once each and not with the "prescience" described by Stanier and van Niel[1]. Sapp[2] elaborates on the more likely origins and evolution of the procaryotic and eucaryotic super-kingdoms found in recent taxonomical structure giving primary credit to Ernst Haeckel[4, 5], Edwind B. Copeland[6], and Herbert F. Copeland[7, 8]. By 1938, Herbert F. Copeland, the son of Edwind B. Copeland, pushed the progression of this evolution toward the birth of four natural phylogenetic kingdoms: Monera, Protista, Plantae, and Animalia[2, 7]. In 1959 the taxonomical model evolved further when Whittaker[9] proposed the addition of a fifth kingdom, *Fungi*, as part of his own four kingdom model that classified Monera as a protist sub-kingdom. Whittaker later refined his model to the widely accepted five kingdom

(Animalia, Plantae, Fungi, Protista, and Monera) taxonomy[10]. Stanier and van Niel's apparent over-exaggeration of Chatton's use of the "procaryote" and "eucaryote" terms[1], though suggested prior to Whittaker's taxonomic model[10], have remained highly popular regardless of their dichotomization of the kingdoms through the separation of Monera as procaryotic and the grouping of the remaining four kingdoms as eucaryotic.

1.1.2 Birth of the Three Domain System

In the early 1960's Carl Woese, a professor of microbiology at the University of Illinois at Urbana-Champaign, began a research program that would attempt to identify the universal phylogenetic tree of life by clarifying the widely unknown phylogeny within the so-called procaryotic super-kingdom[11, 12]. In a recent, more personal account, Woese states his motivation behind establishing the program was to "...restore an evolutionary perspective/spirit to biology" while arguing that the reasoning behind the development of the procaryote/eucaryote classification system was based on a need for formal classification among cellular organisms rather than irrefutable scientific evidence[13]. Acknowledging that proper criteria for the establishment of phylogenetic relationships had not yet been clearly established, Woese began comparing the structures of procaryotic ribosomal ribonucleic acids (rRNA) between species, specifically the 5S rRNAs, by characterizing oligomers produced by specific nuclease digests[11]. This was a necessary approach as complete sequence determination was highly time consuming and the previously used approach of assigning phylogenetic relationships on the basis of protein sequence similarity, as had been done through much of the eucaryotic super-kingdom, was highly ineffective when applied to the bacteria[14].

With the introduction of Fredrick Sanger's macromolecular nucleic acid sequencing technique in 1965[15] (for which Woese personally refers to Sanger as the most important figure in twentieth century biology[16]), Woese and his post-doctoral fellow at the time, George Fox, expanded their approach to include the examination of larger rRNA sequences. By the early 1970's Woese and Fox had managed to tune the Sanger method to fit their requirements[13] and had began studying the sequences of 16S rRNA of a Blue-Green Alga[17], the 16S rRNA of the true bacteria (eubacteria)[18], and the 18S rRNA among the apparent bacterial ancestors within the eucaryotic cytoplasm, referred to as the urkaryotes in phylogenetic analysis[18, 19]. This holistic approach to phylogenetic analysis, based on the evolution of what has long been argued as being the oldest and most evolutionarily stable cellular macromolecule, the ribosome, illustrated two distinct lines of decent between the eubacteria and the ancestral urkaryotes in addition to a third distinct lineage[18]. The identification of this novel lineage lead to the suggestion of a new kingdom within the procaryotic super-kingdom, the *archaebacteria*, which was solely represented by the methanogenic bacteria at the time[18, 20, 21].

Though the original publication was released in the October 1977 issue of the *Proceedings of the National Academy of Science of the United States of America*[21, 22], the first major public appearance of Woese and Fox's novel kingdom was made on November 3, 1977 when major newspapers around the world, including the New York Times, published front page articles quoting Woese's use of the phrase "third form of life."[13, 22]. In a personal account of the events following the public announcement a colleague of Carl Woese, Ralph Wolfe, who was involved with aspects of the discovery and characterization of archaebacterial species, explains how the proposal of the *archaebacteria* as a novel kingdom within the procaryotes was met with substantial resistance.[22] Wolfe recounts the

scientific communities reaction to the news as "...negative with disbelief and much hostility, especially among microbiologists."[22] The strong reaction even included a personal phone call from the Nobel Prize recipient, Salvatore Luria, in which Wolfe was told to "...dissociate [himself] from this nonsense, or... ruin [his] career![22]" Woese, Fox, and their associates would not be discouraged.

In 1980, Fox *et al*[14] further reported the use of 16S rRNA sequence analysis to identify phylogenetic relationships across more than 170 species thus providing the first insight into "procaryotic" phylogeny as a whole. The relationships they observed expanded on the need for the novel archaebacterial kingdom. The phylogenetic analysis revealed each included species fell within one of three primary genealogical lineages[14]. The 16S rRNA sequences within each lineage were nearly identical to one another but differed significantly from those within each of the other two lines of decent. This substantially broadened the spectrum of the new kingdom which now included the methanogens, halophiles, and thermoacidophiles[14].

The identification of the archaebacteria and phylogenetic categorization of the species within this new kingdom eventually culminated a decade later with a proposal by Woese *et al*[23] to establish a novel rank to the existing taxonomic hierarchy. This, they argued, was a necessity as neither the five kingdom taxonomical model, nor the procaryote-eucaryote dichotomy, could accurately describe the evolutionary relationships that had been observed when comparing molecular sequencing data. The new rank, the domain, would supersede the level of kingdom and would hold the formal suffix of *–a*. Woese *et al*[23] specifically stated the domains would be named under the consideration that they would maintain continuity with existing names, would suggest basic characteristics for the overall group, and would avoid any suggestion that the eubacteria and archaebacteria are related to

one another. With this in mind they proposed the formation of the domains *Bacteria*, *Archaea*, and *Eucarya* and suggested the abandonment of their own term "archaebacteria" as it suggests a relationship between the bacterial and archaeal domains of life. Under this system the bacterial domain would consist of the true, or eu-, bacteria and the urkaryotes while the eucaryal domain would consist of all species formerly described as eucaryotic under the procaryotic-eucaryotic dichotomy.

1.2 A Brief Overview of Archaeal Physiological Hallmarks

In order to substantiate the proposal of the three domain system described above, the archaea would have to possess hallmark characteristics that species in neither the bacterial nor eucaryal domains could be found to possess. These would be "genuine phenotypic similarities [that reflect] common evolutionary origin."[24] Among the hallmarks described below are physiological traits possessed by the archaeal domain in its entirety. These include, as outlined by Woese and Fox over the past several decades[13, 14, 18, 23, 24], branched ether-linked lipids, specific structural characteristics of the RNA polymerases, domain specific tRNA base modifications, and composition of the cell wall. Archaea-specific rRNA base modifications are also discussed as these have been found to be highly unique to several archaeal species. Though methanogenesis is not a physiological hallmark of the archaeal domain as a whole and does not apply to the research described in the following chapters, the methanogenic archaeal species are the only organisms of their kind and possess a suite of coenzymes exclusive to their biochemistry. Moreover, the methanogens were integral to the identification of the *Archaea* as a unique domain of life; therefore, methanogenesis has been briefly described.

1.2.1 Archaeal Membrane Lipids

The first fundamental characteristic that differentiates the *Archaea* from the *Bacteria* and *Eucarya* at the domain level lies in the structure of their primary membrane lipids[13]. Many studies over previous years identified "bacterial" species that did not possess the standard ester-linked lipids typical of the *Bacteria*. These species instead possessed diether-linked lipid analogues in which long, branched isoprenoid carbon chains are attached to glycerol via an ether linkage[25-30]. As Woese, Fox, and colleagues[14, 18, 24] continued to build the archaeal phylogenetic tree, diether-linked lipids became an apparent physiological characteristic of the *Archaea*. Synthesis of the isoprenoid side-chains have been shown to occur through head-to-head condensation of isoprene units which in turn can be used in the synthesis of acyclic, monocyclic, or bicyclical diether-linked lipids[31]. The condensation of two diether lipids has been observed to occur in a similar manner to produce a glycerol-dialkyl-glercerol tetra-either lipid product[30]. Although ether-linked lipids have been identified in some eucaryal[30] and bacterial[30, 32] species (and alternatively fatty acid side chains have been observed in some archaeal lipids[30, 33]) to a minor extent, it is the distinct stereochemistry of the carbon at position one of the glycerol backbone of archaeal lipids that set the *Archaea* apart from the *Bacteria* and *Eucarya*[30].

The discovery of ether-linked lipids (prior to the identification of the archaeal domain) lead to the suggestion that they had evolved as an adaptation to the hostile environments in which the "bacterial" species that produced them thrived[26]. The production of an ether-linkage over an ester-linkage would help these organisms adapt to their typically hostile native environments while the presence of branched isoprenoid chains would be an environmental adaptation to elevated temperatures[26]. In more recent years these early suggestions have been confirmed. Lipid profile studies have illustrated changes

in proportion of different ether-linked lipids in the cell membrane under extreme pressure[34] and varying growth temperatures[35]. Further studies revealed cyclization of the isoprenoid side chains yields an increased membrane transition temperature[36] while increased unsaturation of the side chains assists with adaptation to extreme cold[37]. To date, the synthesis of the primary isoprenoids, isopentenyl diphosphate and dimethylallyl diphosphate, and their assembly into chains with subsequent attachment to glycerol-1-phosphate has been thoroughly examined; however, the biosynthetic steps behind the extensive modification (addition of sugar to glycerol; varying levels of desaturation among side chains) remains largely undefined (reviewed by Boucher[30]). Taken together, ether-linked lipids have been shown to be integral to the survival of archaeal species in the hostile environments they tend to flourish in though much remains to be learned.

1.2.2 Archaeal Cell Wall Composition

Another defining physiological characteristic of the *Archaea* described by Woese and Fox[13, 14, 24], with evidence from several studies[38-46], is a strict lack of diaminopimelic and muramic acids (i.e. peptidoglycan) in the outer membrane or cell wall. As our knowledge of the archaeal domain has progressed over the past decades, it has become apparent that the *Archaea* lack a universal cell wall polymer[47]. Archaeal cell walls are now known to be composed of various polymers dependant on the species being observed and can contain lipoprotein[39], sulphated[48], acid[45], or unmodified heteropolysaccharides, glutaminylglycan, methanochondroitin, pseudomurein, proteins, glycoproteins, or glycocalyx[47]. Rather than rigid cell walls archaeal cells tend to possess a proteinaceous exterior surface layer (S-layer) that form regular two-dimensional arrays across the exterior of the cell membrane[47, 49]. Super-structural features and chemical composition of archaeal

S-layers and cell wall polymers are highly variable between genus and species and have recently been reviewed in great detail by Konig *et al*[47].

1.2.3 Archaeal DNA-Dependent RNA Polymerases

Woese[13] has also stated that the characteristics of the DNA-dependent RNA polymerases in the *Archaea*, though they resemble those found in the *Eucarya*[50-53], are a unique physiological characteristic. First described by Zillig *et al*[50] in *Sulfolobus acidocaldarius*, archaeal RNA polymerases are comprised of approximately ten monomeric components including two larger sub-units that resemble the b and b′ subunits of the eubacteria (though the archaeal large sub-units were smaller than any previously described in the eubacteria) in addition to several smaller components[50]. In *Archaea*, however, these sub-units are arranged in a different combination than typical of the eubacteria and the smaller components are analogous to RNA polymerase associated proteins in the *Eucarya*[50]. More recent work has revealed that archaeal RNA polymerases, which are most similar to the eucaryal RNA polymerase II[53], are capable of *in vitro* transcription with only the eucaryal-like transcription factors TBP and TFB[54]. In comparison, the eucaryal RNA polymerase II requires the eucaryal homologs of TBP and TFB in addition to the multimeric transcription factors TFIIE, TFIIF, and TFIIH before *in vitro* transcription will occur[55]. The X-ray crystal structures of two RNA polymerases from the genus *Sulfolobus* have recently been solved[56, 57] and have revealed novel, genus specific structural components[53]. The structure and function of archaeal RNA polymerases[53], and archaeal transcription initiation[52], including these very recent findings, have also been reviewed in detail.

1.2.4 Archaeal Transfer-RNAs

Archaeal transfer-RNAs (tRNA) are another physiological feature Woese[13] has described as defining the third domain. Genes encoding archaeal tRNAs are among the only known archaeal genes to possess introns[58-60] which are processed out of the immature tRNA transcript by a unique endonuclease[61]. Although tRNA processing is a unique physiological feature in archaeal species, it will not be reviewed here as it is thoroughly discussed with respect to considerations that must be made pertaining to introns and Reverse-Transcription Quantitative Polymerase Chain Reactions (RT-qPCR) in Chapter 3. In addition to their method of tRNA processing, however, archaeal tRNAs possess a suite of unique base modifications that have been reviewed by Ramesh Gupta as part of an analysis of tRNAs in the halophilic archaeon *Haloferax volcanii* (previously referred to as *Halobacterium volcanii*)[62]. The most characteristic of these modifications is the presence of an additional pseudouridine (Ψ) residue in the common-arm sequence of each tRNA providing a sequence of Ψ-Ψ-C-G rather than the T-Ψ-C-G sequence commonly observed in eubacterial and eucaryal tRNAs. Some archaeal species, predominately in the kingdom *Crenarchaeota* and the more novel species, *Nanoarchaeum equitans* (Sections 1.3.1 and 1.3.4 respectively), also possess a unique system for processing some tRNAs from the immature tRNA transcript. Several tRNA genes have been identified that are transcribed as two independent immature transcripts then spliced together via a *trans*-splicing mechanism to form the mature tRNA[63]. Although this mechanism is not solely unique to archaeal species[63], the domain does appear to have a large abundance of *trans*-spliced tRNAs.

1.2.5 Archaeal Ribosomal RNA Base Modifications

A number of post-transcriptional modifications have been identified in archaeal rRNAs in the past several decades; however, the specific positions of only a select few have been placed in rRNA sequences. According to the RNA modification Database[64] there have been 15 modifications identified in archaeal rRNA to date, with the 16S rRNA being the most heavily modified (Table 1.1). Levels of modification vary from species to species and have been shown to change in response to cellular stress such as increased culture temperatures[65]. This illustrates the potential for extreme variation in rRNA modification between species as well as within a single species dependant on growth conditions.

Table 1.1. Post-transcriptional base modifications found in the archaeal ribosomal RNAs. The presence (+) and absence (-) of each modification is indicated for each of the 5S, 16S, and 23S rRNAs.

Modification	Abbr.	5S	16S	23S	Reference
N^6-methyladenosine	m^6A	-	+	-	65
2′-*O*-methyladenosine	Am	-	+	+	65
N^6,N^6-dimethyladenosine	$m^6{}_2A$	-	+		21
5-methylcytidine	m^5C	-	+	+	65
2′-*O*-methylcytidine	Cm	+	+	+	65, 66
N^4-acetylcytidine	ac^4C	+	+	+	65, 66
N^4-acetyl-2′-*O*-methylcytidine	ac^4Cm	+	-	-	66
N^2-methylguanosine	m^2G	-	+	-	65
7-methylguanosine	m^7G	-	+	-	67
2′-*O*-methylguanosine	Gm	-	+	+	65
Pseudouridine	Y	-	+	+	68
2′-*O*-methyluridine	Um	-	+	+	65
3-(3-amino-3-carboxypropyl)uridine	acp^3U	-	+	-	69
3-methyluridine	m^3U	-	-	+	65
1-methyladenosine[ab]	m^1A			+	70

[a] *Not Reported in the RNA Modification Database*[64]

[b] *No known reports of modification in 5S or 16S rRNA*

Despite the potential for variability in archaeal post-transcriptional rRNA modifications, most of which are not evolutionarily conserved, the dimethylation of two adjacent adenosines (A1518 and A1519 as per *Escherichia coli* numbering; all modification positions to follow based on this numbering)[69, 71] near the 3′-termini of the 16S rRNA is highly conserved. Currently the only one known archaeal exception to the latter is the crenarchaeotal species, *Sulfolobus solfataricus*, which possesses only a single dimethylation modification in the same region[65]. The equally conserved enzyme[72] responsible for both of these dimethylations is known as KsgA in bacteria and Dim1 in archaea and eucaryotes[72, 73]. Dim1 transfers four methyl groups from four S-adenosylmethionine (SAM) molecules to the two adjacent adenosines converting them to N^6,N^6-dimethyladenosine. It remains unclear as to whether or not the methyl groups are transferred to the adenosines simultaneously or sequentially; however, it is apparent from the KsgA crystal structure[73] that the target adenosines enter the enzyme active site separately and only one molecule of SAM is bound at a time[72] which suggests sequential transfers are likely.

A second modification of interest is the previously described 3-(3-amino-3-carboxypropyl) uridine (acp^3U)[69, 74] which has been shown to occur in the 16S rRNA of the halophilic archaeon, *Haloferax volcanii* at position 966. This unique modification was thought to exist in tRNA only and had never previously been observed in rRNA sequences[69]. The *H. volcanii* acp^3U modification corresponds to a universally conserved modification site, which in bacteria contains an m^2G modification directly adjacent to an m^5C modification at position 967. The same site in eucaryotes is typically an evolutionarily conserved uridine residue which is commonly modified to 1-methyl-3-(3-amino-3-carboxypropyl) pseudouridine; a structural relative to the acp^3U modification[69].

Previous studies have shown that the 3-amino-3-carboxypropyl group is transferred from SAM to uridine when synthesized in tRNAs[75]; however, the mechanism behind this transfer and the function of the resulting modification is apparently yet to be determined.

1.2.6 Methanogenesis

The methanogens have long been regarded as an ancient lineage of organisms that originally adapted to an early anaerobic earth[20, 21]. Although this would appear to be entirely plausible, more recent phylogenetic studies have suggested that the methanogens diverged later than previously anticipated as the Thaumarchaeotal lineage[76] apparently diverged first (Section 1.3.5). Nonetheless, as the first recognized archaeal organisms, the methanogenic archaea comprise several orders within the euryarchaeotal kingdom and possess a highly unique biochemistry. The well studied production of methane is the primary energy yielding metabolism of the methanogenic archaea and has been thoroughly reviewed by Ferry and Kastead[77]. As described in this review, methanogens utilize several cofactors, many of which can be found in the other domains of life while others are specific to the methanogenic archaea. The first methanogen-specific cofactor, 2-mercapto-ethanesulfonic acid (Coenzyme M), was described by Taylor and Wolfe in 1974[78] and is the smallest known enzymatic cofactor currently known[77]. Additional cofactors include methanofuran, 7-mercaptoheptanoyl-threonine phosphate (Coenzyme B), methanophenazine, 5,6,7,8-tetrahydromethanopterin, 5-hydroxybenz-imidazolylcobamide (Factor III), $Factor_{430}$, Coenzyme F_{420}, and molybdopterin guanine dinucleotide[77]. The biological production of methane typically occurs via one of two primary metabolic pathways which involve the fermentation of acetate or the reduction of carbon dioxide[77]. In addition to the primary pathways, species within the orders *Methanosarcina* and

Methanoccoides are capable of producing methane through the conversion of single carbon compounds such as methanol or methylamines[77].

One of the most interesting aspects of methanogenesis has been made apparent through investigations into the molecular biology of methane production. These investigations have lead to the discovery of the novel amino acid, pyrrolysine[79] (Figure 1.1) which has been found to be encoded by a UAG codon and is used to charge a specific tRNA directly via the enzyme pyrrolysyl-tRNA synthetase[80, 81]. In 2006, Mahapatra *et al*[82] conducted a study in which they knocked out the gene encoding the dedicated pyrrolysine tRNA from *Methanosarcina acetivorans*, a species capable of utilizing several substrates for methane production and carbon assimilation including various methylamines. Mahapatra *et al*[82] recognized that growth of their knockout strain on substrates that required the use of proteins or processes requiring pyrrolysine should be lethal to the organism. With this in mind, growth was assessed on media containing the typical *M. acetivorans* substrates: methanol, acetate, monomethylamine, and dimethylamine. The knockout strain was capable of growth on methanol and acetate over extended growth periods of several years and these substrates were interchangeable; however, knockout strain growth on mono- or dimethylamine was not observed even after extended six month incubation periods[82]. Additionally, unlike the wild-type strain, the *M. acetivorans* knockout strain was apparently unable to assimilate nitrogen from mono- or dimethylamines while starved of other nitrogen containing substrates[82]. Moreover, methyl transfer from methylamines to Coenzyme M was not observed in the knockout strain[82]. Taken together, this evidence suggests pyrrolysine is required for a *Methanosarcina*-specific biochemical process that transfers a methyl group from methylamines to Coenzyme M and may play a role in nitrogen liberation from these compounds under limiting conditions.

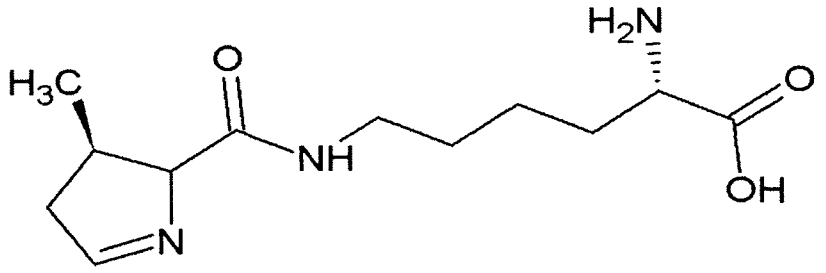

Figure 1.1. Pyrrolysine identified in methylamine methyltransferases of the order *Methanosarcina*. The naturally occurring amino acid is used to charge a dedicated tRNA via pyrrolysyl-tRNA synthetase.

1.3 Phylogenetic Structure of the Archaeal Domain and its Kingdoms

In addition to Woese and Fox's proposal for the three domain system[23], they also suggested the formation of the novel kingdoms *Crenarchaeota* and *Euryarchaeota* within the archaeal domain. The euryarchaeotes were to include the three methanogenic lineages, the extreme halophiles, the sulfate-reducing species and the thermophiles of the genus *Thermoplasma* and the *Thermococcus/Pyroccocus* group. Alternatively, the crenarchaeotes would include the thermoacidophiles, the extreme thermophiles, and the sulphur-dependant species[23]. Today five kingdoms (*Crenarchaeota*[23], *Euryarchaeota*[23], *Korarchaeota*[83], *Nanoarchaeota*[84], and *Thaumarchaeota*[76]) have been identified encompassing more than 110 individual archaeal species for which whole genomes have been sequenced and annotated[85, 86]. Several hundred additional species have been identified[87], whether it be through molecular ecology studies involving analysis of environmental samples or isolation via pure culture. Presumably owing to the lack of the higher-level taxonomic grouping of phylum, and the lesser number of species in each, relative to their bacterial and eucaryal counterparts, the archaeal kingdoms are commonly referred to as phyla throughout the literature; however, because they were originally proposed as such by Woese[23], they will

continue to be referred to as kingdoms below. Each of the five kingdoms is briefly outlined below including an introduction to a few species of particular interest.

1.3.1 *Crenarchaeota*

The kingdom *Crenarchaeota* was aptly named from the Greek word '*crenos*', meaning spring or origin[23, 76], which was given for the believed physiological resemblance of the species to the ancestor of the domain considering their tendency towards hyperthermophilicity[23]. The kingdom is currently comprised of the orders *Sulfolobales*, *Acidilobales*, *Desulfurococcales*, and *Thermoproteales*[76, 85, 86] from which 35 whole genomes have been sequenced[76, 85, 86]. The kingdom is home to many highly unique species that could each be discussed at length; however, for the purposes of this introduction, only the genera *Sulfolobus* and *Ignicoccus* are included here.

The genus *Sulfolobus* is comprised of species that have arguably been the most studied of the archaeal domain. The genus was first proposed in 1972 as a novel bacterial taxonomic group[42] and was reclassified as one of the original archaeal genera when the archaeal domain was proposed by Woese and Fox[14]. Species within this genus are described as typically spherical cells that utilize sulphur or other simple organic compounds as primary growth substrates[42]. These species are also characterized as thermophilic acidophiles with optimal growth typically occurring at temperatures of 70-75°C and at pHs between 2 and 3[42]. Both *Sulfolobus acidocaldarius* and *Sulfolobus solfataricus* can easily be referred to models for the kingdom as these species have been widely used as the "go-to" organisms for investigations involving the *Crenarchaea* and have become instrumental in advancing our knowledge of the crenarchaeotal kingdom.

One of the most interesting groups of organisms within the *Crenarchaea* are those comprising the genus, *Ignicoccus*. These are arguably the most unique of the entire archaeal domain in terms of cellular morphology as they are the only known archaeal species that do not possess either a proteinaceous S-layer or any other cell wall polymer[47], but rather a cytoplasmic membrane, variable periplasmic space, and an outer sheath resembling those of gram-negative bacteria[49, 88]. Members of this genus are the only species described as obligate chemolithotrophic sulfate reducing organisms within the order *Desulfuro-coccales*[88]. Cells grow either singly or in pairs and possess several flagella-like appendages[88]. Additionally, at least one *Ignicoccus* species, *Ignicoccus hospitalis*[88], has been shown to grow in the presence of a parasitic archaeal symbiont[84] (Section 1.3.4).

1.3.2 *Euryarchaeota*

The kingdom *Euryarchaeota*, was named from the Greek word '*euryos*', meaning diverse[76], broad, wide, or spacious[23], which was given for the spectrum of ecological niches these species fill and for the array of metabolisms they encompass[23, 76]. The kingdom is currently comprised of the orders *Archaeoglobales*, *Halobacteriales*, *Methanobacteriales*, *Methan-ocellales*, *Methanococcales*, *Methanomicrobiales*, *Methanopyrales*, *Methano-sarcinales*, *Thermococcales*, and *Thermoplasmatales* from which 75 genomes have been sequenced[76, 85, 86]. As with the *Crenarchaeota*, *Euryarchaeota* is home to several species of interest; however, not all of these can be discussed here. Therefore, in addition to the well studied methanogenic species in Section 1.2.6, the discussion of the euryarchaeotic species below will include two species of personal interest: *Ferroplasma acidarmanus*, chosen for the intriguing environment in which it thrives, and *Haloquadratum walsbyi*, chosen for its highly unique morphological features. The halophilic archaeon, *Haloarcula marismortui*, is

discussed in greater detail below (Section 1.4) as it is the subject of interest in the investigation detailed in the chapters to follow.

The acidophilic archaeon, *Ferroplasma acidarmanus* strain fer1, was first isolated from within the biofilms of the highly acidic, metal rich waters of the abandoned Iron Mountain mine in northern California[89]. In 2003 two additional strains, MT16 and MT17 were isolated from a biooxidation pilot plant[90] of which only the MT17 strain has been well characterized. *F. acidarmanus* growth can be observed across a pH range of 0.0-2.5 with optimal growth occurring between pH 0.6 and 1.4[89, 91]. Both the fer1 and MT17 strains grow heterotrophically on yeast extract[89-91]. The two strains differ in that MT17 has an optimal growth temperature of 39°C[90] whereas fer1 grows optimally at 42°C[89, 92]. Both strains will oxidize ferrous iron and growth will occur in acidic media supplemented with pyrite (FeS_2) sediments and 0.02% yeast extract[89, 90]. This oxidation of pyrite results in acid production via the reaction $FeS_2 + 14\ Fe^{3+}\ 8\ H_2O \rightarrow 15\ Fe^{2+} + 2\ SO_4^{2-} + 16\ H^+$ [93] and *F. acidarmanus* mediated pyrite dissolution such as this is thought to contribute to the acidic environments where these organisms are commonly located[92]. Iron oxidation of this nature has also been hypothesized to be a major source of microbial energy generation in low pH environments[92].

The second species of interest is the halophilic archaeon, *Haloquadratum walsbyi*. This morphologically unique organism was first described by A.E. Walsby in 1980 in water samples taken from hyper-saline pools near the Red Sea[94]; however, the species was not successfully isolated in pure culture until the middle of the last decade[95, 96]. These cells grow in a distinct, square morphology with dimensions of 2 x 2 x 0.2μm and have been observed in the form of aggregate sheets up to 40 x 40μm[95, 96]. The organism also contains

distinct gas vesicles; a morphological feature that Walsby used to first identify the species as a micro-organism in 1980[94]. Long referred to as "Walsby's Square Haloarchaeon", *Hal. walsbyi* was given its current name in 2007 after 16S rRNA sequencing confirmed it as a novel species of the family *Halobacteriaceae*.

1.3.3 *Korarchaeota*

The kingdom *Korarchaeota* was proposed in 1996 by Barns *et al*[83] after molecular ecology studies revealed the presence of several species of a distinct archaeal lineage in environmental samples obtained from hot springs in Yellowstone National Park. The novel lineage was named from the Greek words "*koros*", meaning young man, or "*kore*", meaning young woman, which was given to describe the apparent early divergence of these organisms from the larger crenarchaeotal and euryarchaeotal kingdoms[83]. It was immediately recognized that the *Korarchaeota* would not be able to be confirmed as a novel lineage without additional studies and, preferably, with the cultivation of a korarchaeotal species in pure culture[83]. As of 2006 a pure korarchaeotal culture had not net been obtained; however, one of the strains first described by Barns *et al*[83] has been highly enriched[97]. To the best of our knowledge, no single species has been isolated in pure culture to date. All known korarchaeotal strains have been identified via additional molecular ecology studies utilizing *Korarchaeota* specific primers for the 16S rRNA[98, 99]. To date, a single genome has been sequenced using the enriched cultures described above[97]. This species, *Candidatus Korarchaeum cryptophilum*[97] is currently the only confirmed organism within the kingdom though evidence suggests many more exist.

1.3.4 *Nanoarchaeota*

The kingdom *Nanoarchaeota* (the dwarf archaea) was first proposed in 2002 with the discovery of a novel archaeal species that could only be co-cultured[84] with the hyperthermophilic crenarchaeotal species, *Ignicoccus* sp. strain KIN4I[84, 100] (now known as *Ignicoccus hospitalis*[88]), which as described above, belongs to a genus that is highly unique in its own right. The new species, which thrives as a parasitic symbiont of *Ignicoccus*, was named *Nanoarcheaum equitans*, meaning "riding the fire sphere[84]. *N. equitans* is currently the only species representing *Nanoarchaeota* and is one of the smallest coccoid organisms currently known with a cell diameter of 400nm[84] and a genome of just 480Kb[101, 102]. The size of the genome, which lacks genes for various critical metabolic pathways (i.e. lipid synthesis[103]) emphasizes the reliance of *N. equitans* on its *Ignicoccus* host[47, 101]. This has been further confirmed after analysis of lipids from both *N. equitans* and *Ignicoccus* sp. KIN4I suggests *N. equitans* does not synthesize its own membrane lipids, but rather obtains lipids via selective uptake from its host[103]. The small genome also codes for several half-tRNA genes which are transcribed separately before being spliced together via the highly unique *trans*-splicing mechanism[104-106] described above (Section 1.2.4). With the publication of some phylogenetic studies that have suggested *Nanoarchaeota* is better grouped as a lower-level taxonomic rank within the *Euryarchaeota* when considering an elevated rate of evolution[76, 102], the validity of this kingdom is still being debated. The kingdom *Nanoarchaeota* and its lone occupant have been further reviewed by Huber *et al*[107].

1.3.5 *Thaumarchaeota*

From the Greek word '*thaumas*', meaning wonder[76], *Thaumarchaeota* is the newest of the five archaeal kingdoms. The formation of the novel kingdom was proposed in 2008 by Brochier-Armanet *et al*[76] based on phylogenetic studies that used ribosomal proteins and small subunit rRNA sequences in addition to genomic core gene sequences common to either *Euryarchaeota*, *Crenarchaeota*, or both. These studies showed the mesophilic, ammonia-oxidizing species, *Cenarchaeum symbiosum*, which was previously believed to belong to the kingdom *Crenarchaeota*, actually branched at a node in the phylogenetic tree that placed the organism in a sister group to a clade comprised of the *Euryarchaeota* and *Crenarchaeota* kingdoms[76]. This would suggest the *Thaumarchaeota* speciated before the speciation of the two larger kingdoms[76]; however, the use of a single species to properly identify a novel kingdom was less than desirable. Several molecular ecology studies have identified a substantial number of potential thaumarchaeotal genetic signatures in a wide variety of mesophilic environmental samples (briefly reviewed by Spang *et al*[87]). The large number of detected ammonia-oxidizing archaeal species has led to suggestions that these organisms may play a substantial role in global nitrogen and carbon cycles[76, 87]. The isolation[108] and subsequent genome sequencing[109] of the marine archaeon *Nitrosopumilus maritimus*, followed by the isolation[110] and draft genome sequence determination[87] of *Nitrososphaera gargensis* has helped to resolve the phylogenetic relationships within the proposed *Thaumarchaeota* kingdom. Additional phylogenetic analysis including these novel species has shown that the *Thaumarchaeota* are in fact a distinct, and deeply branching, lineage within the archaeal domain; however, the phylogenetic positioning of the thaumarchaeal kingdom can not yet be unambiguously resolved.

1.3.6 Evidence for a Sixth, Un-named Archaeal Kingdom

A recent study utilizing the composite genome sequence of a currently uncultured archaeon, *Candidatus Caldiarchaeum subterraneum*, has suggested the potential for an additional novel kingdom[111]. Nunoura *et al*[111] conducted phylogenetic studies in 2011 that suggested the uncultured species belonged to a novel taxonomic group that branched deeply within the archaeal domain. The genomic sequence was shown to contain commonalities with the other archaeal kingdoms as well as some features typical of the *Eucarya*[111]. Unfortunately, *C. subterraneum* is the only species currently known to exist in this phylogenetic group and it is not yet available as a pure culture. Until additional species related to this organism can be identified via molecular ecology studies or isolation in pure culture and included in new phylogenetic analyses, the existence of this potentially novel archaeal kingdom can not yet be confirmed.

1.4 A Brief Introduction to *Haloarcula marismortui*

1.4.1 Isolation and Growth of the Laboratory Strain

The current laboratory strain of *Haloarcula marismortui* was isolated by Ginzburg *et al*[112] during the 1960's under the name "*Halobacterium* of the Dead Sea". The organism was believed to have been previously identified as *Halobacterium marismortui* by Elzari-Volcani during his work at the Dead Sea[113]; however, the strain was never deposited into a culture collection and was eventually considered lost[114]. The new strain isolated by Ginzburg *et al*[112] was eventually deposited into a culture collection and was later described by Oren *et al*[114] who identified the organism as being morphologically similar to *Haloarcula californiae* and as having similar properties to those first described by Elzari-Volcani for *Halobacterium marismortui*; however, the new organism possessed properties

that were exceptions to those described for both of the latter species leading to the proposal of the name *Haloarcula marismortui* (marismortui meaning "of the Dead Sea").

Har. marismortui is a motile[115] (though originally described as non-motile[114]), obligatorily aerobic, chemoorganotroph that has the ability to utilize a wide variety of compounds such as glucose, sucrose, fructose, glycerol, acetate, succinate, and malate as sole carbon sources[114]. Cells have been described as pleomorphic, flat disks with measurements of 1-2 by 2-3μm[114] with a tendency toward the rod shape[112]. Growth occurs in media containing 1.7 to 5.1M sodium chloride (NaCl) with optimal growth occurring in concentrations of 3.4 to 3.9M at an optimal temperature of 40 to 50°C[114]. The intracellular potassium ion concentrations within *Har. marismortui* have been reported to be in excess of 3M which may be explained by previous reports suggesting that many halophilic archaeal species sequester potassium ions within the cell[116]. It is believed *Har. marismortui* sequesters the ion and achieves this substantially elevated internal salt concentration as a mechanism of balancing the osmotic pressure experienced in its native hyper-saline environment[112].

1.4.2 *Haloarcula marismortui* as a Model Organism

Haloarcula marismortui is arguably one of the most important organisms to modern molecular biology as it is widely regarded as a model for studies pertaining to the eucaryotic ribosome. Attempts to assess the structure of the highly complex, large ribosomal sub-unit began in the 1980s with reports of the successful isolation of ribosome crystals obtained using ribosomes from *Har. marismortui*[117]. The native environment in which these cells thrived suggested their proteins and nucleic acids should be highly stable in high-salt environments, thus making their ribosomes a favourable choice for attempts at

crystallization using solutions near saturation. As time progressed, the laboratory of Thomas A. Steitz also began working toward obtaining a high resolution X-ray crystal structure of the large ribosomal sub-unit and in 1998 the group published a crystal structure with 9Å resolution[118]; far exceeding the previous resolution of only 20Å[119]. Steitz's group followed this up a year later with the release of a structure with 5Å resolution[120] then, at the turn of the decade, they published the current 2.4Å resolution crystal structure of the large ribosomal sub-unit[121]. Steitz was awarded a Nobel Prize for this work in 2009[119] and today the crystal structures obtained from *Har. marismortui* are widely used as structural models for the eucaryotic ribosome due to their simplistic similarities.

1.4.3 Novel Photo-Active Rhodopsins

Most halophilic archaeal species possess several microbial rhodopsins comprising a family of trans-membrane proteins that utilize a photo-response to mediate ion transport as a means of harvesting solar energy or act as receptors for phototaxis[122, 123]. Until recently, haloarchaeal species were believed to possess 4 unique rhodopsins[124, 125]. These consisted of a photo-dependent proton pump (bacteriorhodopsin)[126], a photo-dependent chloride pump (halorhodopsin)[127, 128], and two sensory rhodopsins, one of which mediates both attractant and repellent phototaxis (sensory rhodopsin I)[129, 130] and one of which mediates repellent-specific phototaxis (sensory rhodopsin II)[123, 131]. A recent study by Fu *et al*[123] has shown *Har. marismortui* to be unique from other halophiles in that it possess a novel, highly refined six-rhodopsin system. This system provides *Har. marismortui* with a more diverse photo-sensing ability in comparison to other halophiles as the absorption spectra of the six rhodopsins provides a broader distribution of wavelength maxima utilized. A photo-driven ion transport system that contains two isochromatic rhodopsins that, as of yet,

appear to be identical in function and level of expression is also present. Additionally, Fu *et al*[123] report a sensory-like rhodopsin protein containing a shortened cytoplasmic region of unknown function that does not induce phototaxis and suggest this novel photo-dependent protein may play a role in photo-adaptation or regulation of a circadian rhythm.

1.5 Objectives

Today our expanding knowledge base of these unique and fascinating organisms is made clearly evident by searching the online PubMed Central archival database (http://www.ncbi.nlm.nih.gov/pmc/) using the search term "*Archaeal*" which returns 11,000 related articles published since January of 1995. A brief survey of these articles suggests very few, if any, studies have been conducted pertaining to the cellular response of halophilic species to changes in extracellular ion concentrations. More specifically, there appears to be only a single publication that has investigated intracellular potassium ion concentrations in haloarchaea and this publication provided evidence that has contradicted recent, unpublished findings in our lab. Additionally, of the studies utilizing RT-qPCR techniques as a primary tool for quantitative molecular studies in archaeal species, none appear to adhere to the guidelines for the minimal information for the publication of quantitative PCR experiments (the MIQE Guidelines; Chapter 3) as published in 2009[132]. A concerning lack of confirmed RT-qPCR reference genes have been identified in archaeal species which is likely the largest contributing factor to the latter. In an attempt to address both of these concerns, the goal of the research described in the following chapters is two-fold:

1. Assess the cellular response to potassium stress in *Haloarcula marismortui* via determination of cellular generation times and changes in intracellular potassium concentration across a wide range of extracellular potassium concentrations; and,

2. Identify multiple candidate RT-qPCR reference genes and confirm the stable expression for each across various growth conditions in order to provide the archaeal scientific community with a starting point for reference gene identification in other species.

Changes in cellular generation time is a key indicator of cellular stress and will illustrate *Har. marismortui*'s ability to adapt to changes in its native environment. The additional assessment of the intracellular ion concentrations obtained from cells grown at various extracellular potassium concentrations will not only confirm or deny previously reported intracellular concentrations, but will provide valuable information pertaining to physiological changes that occur while *Har. marismortui* adapts to its changing environment. Moreover, this could potentially provide insight into changes in ion transporter activity and physiologically optimal ion concentrations. In addition to the latter, the identification of novel RT-qPCR reference genes will provide the framework for further studies regarding differential gene expression in *Har. marismortui*. More specifically, this will provide the framework for an RT-qPCR-based investigation into the differential expression of membrane bound ion transport proteins and further define the first of the two above objectives.

Chapter Two

Evaluation of Potassium Stress Responses in the Halophilic Archaeon, *Haloarcula marismortui.*

2.1 Introduction

2.1.1 Growth Characteristics of Halophilic Species

Growth characteristics of halophilic archaeal species have been only minimally studied. These studies have a tendency toward changes in growth characteristics with varying temperature and, to the best of our knowledge, have not examined changes in growth characteristics in response to specific ion stresses. A more recent investigation of haloarchaeal growth kinetics by Robinson *et al*[133] utilized a variety of species of the family *Halobacteriaceae* and found cellular generation times within the family vary from 1.5 to 3 hours and identified several species that have multiple temperature optima. The species most closely related to *Haloarcula marismortui* that was examined in this study was *Haloarcula vallismortis*, which produced an optimal cellular generation time of 3.04 ± 0.20 hours at an optimal temperature of 43 - 49°C[133].

The extent to which *Har. marismortui* growth has been characterized is limited largely to the initial characterization that accompanied the proposal of the organism as a novel species in 1990[114]. The focus of the other studies pertaining to *Har. marismortui* growth maintain a focus of examining substrate utilization and these publications appear to be extremely rare. A survey of the literature has so far revealed only a single publication by Ginzburg *et al*[112] stating a cellular generation time of 5-6 hours for an unknown Dead Sea isolate, though other such publications may exist. This isolate is described as a highly pleomorphic cell type that demonstrated a tendency toward the rod shape and produced pink coloured cultures that deepened in colour with age[112]. Though this description does not

allow for a definitive identification of the isolate as *Har. marismortui*, it does describe this particular species extremely well. Several more recent publications have since cited this report as referring to *Har. marismortui*.

2.1.2 Ion Transport in *Haloarcula marismortui* and Related Species

2.1.2.1 Mechanisms of Osmoregulation

Ion transport in *Har. marismortui* has been studied to a far greater extent than the organisms growth characteristics. Ginzburg *et al*[112] produced what is likely the first report of ion transport in this species when they determined the cell water volume (1.22 ± 0.02mm^3; 1.22 ± 0.02μL)[112] and corresponding concentrations of K^+, Na^+, and Cl^- within an unknown halophile isolate from the Dead Sea (described above) using primarily gravimetric methodologies. This investigation examined the intracellular concentrations across various stages of growth and reported intracellular ion concentrations ranging from 3.7 - 5M K^+, 0.5 - 3M Na^+, and 2.3 - 4.2M Cl^-. Ginzburg *et al*[112] acknowledged that these concentrations were extreme and solutions of 4-5M KCl and 1-3M NaCl can not be prepared due to limitations of solubility. They then go on to suggest that the activity of the potassium ion is somehow limited within the cells, which would allow for extremely elevated concentrations. Moreover, the permeability of the membrane[134] lead them to suggest the mobility of K^+ is largely restricted within the cells.

Neutron scattering has been recently used to examine cell water movement in *Har. marismortui* and revealed a slow-moving water component that accounts for approximately 76% of the total cell water[135]. It has been suggested that this slow-moving cell water component is arranged around proteins within the cell as a solvation shell that may bind large amounts of K^+ in order to maintain the viability of halophilic proteins. If this is the

case, it would confirm the notion made by Ginzburg *et al*[112] regarding restricted K^+ mobility.

As explained in an excellent review by Oren[136] just over a decade ago, microorganisms in all domains of life utilize one of two primary mechanisms for survival: (i) Cells sequester salts internally to concentrations that are osmotically equivalent to the extracellular salt concentrations forcing intracellular systems to adapt to, and function in, a hyper-saline environment, or; (ii) Cells maintain a low-salt intracellular environment and balance osmotic pressure through the synthesis of organic solutes, such as glycerol, glycine betaine, or sucrose, eliminating the need for the adaptation of intracellular systems. By these definitions, the observations described in the paragraphs above obviously indicate *Har. marismortui*, along with all other halophiles of the order *Halobacteriales*[136], utilizes the first of these mechanisms for survival in their native hyper-saline environment.

2.1.2.2 Maintenance of a Potassium Gradient

Members of the order *Halobacteriales* utilize the proton electrochemical gradient across the cell membrane to drive the expulsion of Na^+ and sequestration of K^+ [136]. As reviewed by Oren[136], this gradient is maintained via respiratory electron transport during aerobic growth or substrate-level phosphorylation through membrane ATPases. In the case of members of the family *Halobacteriaceae*, which includes *Har. marismortui*, the proton gradient is generated directly via the photosensitive proton pump, bacteriorhodopsin[126, 136]. The established proton gradient, as maintained by any of the latter mechanisms, is then used in conjunction with Na^+/H^+ antiporters as a primary mechanism to drive and maintain the Na^+ gradient across the cell membrane[136]. Additionally, the accumulation of Cl^- has been shown to occur via the photosensitive halorhodopsin transporter[127, 128] and is also

believed to occur through a co-transport mechanism with Na^+ movement back into the cell[136].

The accumulation of K^+ has been argued to occur via passive diffusion through a uniport system allowing for accumulation proportional to the magnitude of the electrochemical potential across the cell membrane as per the Nernst equation[136, 137]. A study of K^+ transport in the Haloarchaeon, *Haloferax volcanii*, has shown, however, that the intracellular concentrations of K^+ observed in this organism can not be accounted for by passive processes alone and ATP hydrolysis is required to actively transport K^+ into the cell in order to reach the 3.6M intracellular concentrations that are maintained by *Hfx. volcanii*[136, 137]. An ATP-regulated, low-affinity K^+ transporter that is similar to the Trk system found in *Escherichia coli* has also been documented[136, 137]. Overall, the accumulation of intracellular K^+ as a mechanism of osmoregulation is more energetically favourable than the mechanism of synthesizing or sequestering organic solutes ($ATP:K^+$ costs reviewed by Oren[136]); however, a trade off for an adaptation of cellular processes to excessive salt concentrations is required.

2.1.3 Study Specific Objectives

The primary aim of this study is to assess the cellular responses to potassium stress in *Haloarcula marismortui* as a means of addressing the lack of growth characterization for this species. In order to achieve this goal cellular generation times need to be properly assessed. Growth was evaluated across a variety of conditions that encompass extreme changes in extracellular potassium concentration. Growth in the presence of the alternative monovalent cations of lithium, rubidium, and caesium was also evaluated. Additionally, the intracellular concentrations of K^+, Li^+, Rb^+, and Cs^+ were determined using modern

analytical techniques conducted on cells grown under the extreme extracellular potassium concentrations used for cellular generation time analysis. This was done in an attempt to gain insight into *Har. marismortui*'s ability to cope with ion concentration changes in its native environment and provide valuable insight into the organisms ion transport capabilities. These capabilities were further assessed via the determination of intracellular alternative ion concentrations when potassium is at a minimal extracellular concentration and the alternative ion concentration is elevated. It is our hope that this evaluation of growth and subsequent determination of intracellular ion concentrations in the presence of alternative monovalent ions may prove valuable while refining the known mechanisms of halophilic osmoregulation.

2.2 Methods

2.2.1 Preparation of *Haloarcula marismortui* Cell Cultures

Har. marismortui cells were grown in 23% Salt Water Modified Growth Media (23% S.W. MGM) as described by Rodriguez-Valera *et al*[138, 139]. Cultures were incubated at 45°C with constant shaking at 250rpm. Consistent lighting conditions were maintained throughout growth to eliminate the possibility of changes in cellular generation times due to changes in stimulation of photosensitive membrane proteins (bacteriorhodopsin[126], halorhodopsin[127, 128], etc). Cells were continuously sub-cultured at mid-exponential growth (OD_{600} = 0.4-0.6) as a means of maintaining continuously doubling cultures. Once cultures had been sub-cultured no less than 3 times, cells were defined as being in "balanced growth". Media containing monovalent ions alternative to potassium was prepared as per the methods out lined by Rodríguez-Valera[138, 139] with modifications. High-purity NaCl (Fluka; 99.999% by trace metal analysis) was used in the preparation of the initial 30% salt-

water solution and KCl was excluded. Yeast extract and Peptone were added as described and 3.5M salt solutions containing the desired ionic salt (LiCl, RbCl or CsCl) were added with a volume that provided a 100mM final concentration. Media was then brought to a pH of 7.5 using Tris-base and the media was adjusted to the final volume using Milli-Q H_2O. In the case of the 100mM LiCl growth media, the LiCl salt was added before the media was sterile filtered and autoclaved. For the RbCl and CsCl medias, the media was autoclaved at a reduced volume before adding the 3.5M salt solutions, adjustment of pH and volume, and sterile filtering. Trace quantities of potassium in the high-purity NaCl stock leaves a K^+ concentration in these media conservatively estimated at an 8mM minima.

2.2.2 Determination of Cellular Generation Times

Cells were grown to balanced growth as described above in standard 23% S.W. MGM as well as in media containing 8, 20, 220, 520, and 720mM KCl. Potassium concentrations below 8mM were not obtainable due to trace levels of potassium in the high-purity NaCl salts used. Cultures were grown as a biological triplicate of technical replicates (i.e. one triplicate of cultures per biological replicate). Cell densities for each culture were measured via spectrophotometry at a wavelength of 600nm (OD_{600}). Growth curves were constructed by measuring cell density at least once per generation time as determined by an experimental test culture. Measurements were averaged across all replicates then plotted against the growth time. In order to determine cellular generation times (Figure 2.1) the data were fit to the exponential portion of the curve using the equation:

$$A = A_0 e^{kt}$$

Where A = OD_{600}, A_0 = Initial OD_{600}, k = growth constant, and t = cellular generation time.

2.2.3 Determination of Cell Density

The cell density of a pink, rod shaped Halobacterium that had been isolated from the Dead Sea has been previously reported as 1.20g/mL[112]. As this description fits that of *Har. marismortui*, the technique used in this report was adapted and repeated to confirm this result. Solutions of NaCl and sucrose were prepared with a final NaCl concentration of 3.5M and variable sucrose concentrations to vary densities. Cells at mid-exponential growth (OD_{600} = 0.4-0.6) and at saturation (OD_{600} = >0.7) were pelleted independently by centrifugation. Pellets were resuspended in a small volume of 23% S.W. MGM to create a high viscosity cell suspension that was layered onto the NaCl/Sucrose solution. The layered suspensions were then centrifuged at 6,000rpm for 30 seconds. The sucrose solution that, upon visual inspection, allowed 50% of the high viscosity suspension to travel 50% of the height of the solution within the microcentrifuge tube was considered to be of equal density to the cells. Density of NaCl/Sucrose solutions was determined gravimetrically.

2.2.4 Determination of Average Cell Volume

Using the procedures for constructing a growth curve, optical densities were measured at least once every experimentally determined generation time. Cells were then diluted 50-1000 fold and counted using a standard haemocytometer. A standard curve plotting the number of cells/mL against the optical density of a cell culture grown under standard conditions was constructed and a line of best fit applied (R^2=0.997; Figure 2.2). To determine the average mass of a single cell, the number of cells in a given volume of culture was determined using the equation $y = (2\times10^9)x - (2\times10^7)$ where y is the cellular density in terms of number of cells/mL of culture and x is the optical density of the culture. The volume of culture was centrifuged at 13,000rpm for 5 minutes and the media was

removed completely by pipette. The pellet was centrifuged a second time and residual media was aspirated off of the pellet. A thin residue was observed on the walls on the micro-centrifuge tube containing the pellet after aspiration so a base-line mass for this residue was obtained by aspirating an equivalent volume of growth media independently. This process was completed three times using 500, 750, and 1250mL of cell culture. The average cellular density was then determined using the obtained average cellular mass and the density equation (ρ=m/v).

2.2.5 Evaluation of Intracellular Ion Concentrations

Cells in balanced growth in media containing 8mM KCl, 100mM KCl (23% S.W. MGM) 720mM KCl, 100mM LiCl, 100mM RbCl, and 100mM CsCl were centrifuged at 13,000rpm for 5 minutes and all media was removed via pipette. Cell pellets were washed twice using the 8mM KCl media as a means of washing excess ion away without lysing cells. Pellets were centrifuged a second time and residual media was removed via pipette. Cells were lysed in 5mL 1% nitric acid (HNO_3; Milli-Q H_2O) and sonicated to break up small particles in the lysate. Induction coupled plasma-mass spectroscopy (ICP-MS; conducted by University of Northern British Columbia's Central Equipment Laboratory) was used to determine the concentration of K^+, Li^+ Rb^+, and Cs^+ within each cell lysate. The measured concentrations were used to determine the approximate quantity of ion within a single cell using the standard curve equation for cell number/mL described above, which in turn was used to determine the intracellular concentration for each ion using the previously determined average cellular volume.

2.3 Results

Haloarcula marismortui growth was observed across the broad range of extracellular potassium concentrations examined. Generation time assessment confirms the use of our standard growth conditions (100mM KCl) provide an optimal K^+ concentration and produced a generation time of 4.19 ± 0.14 hours. Growth occurs at concentrations as low as 8mM KCl and as high as 720mM KCl. The cellular generation time remained highly stable across the elevated K^+ concentrations examined; however, as the extracellular K^+ concentration was decreased the generation time increased sharply. Generation times for *Har. marismortui* growth on alternative monovalent ions was not assessed as growth was observed to become increasingly slower after each sequential sub-culture (Section 2.4).

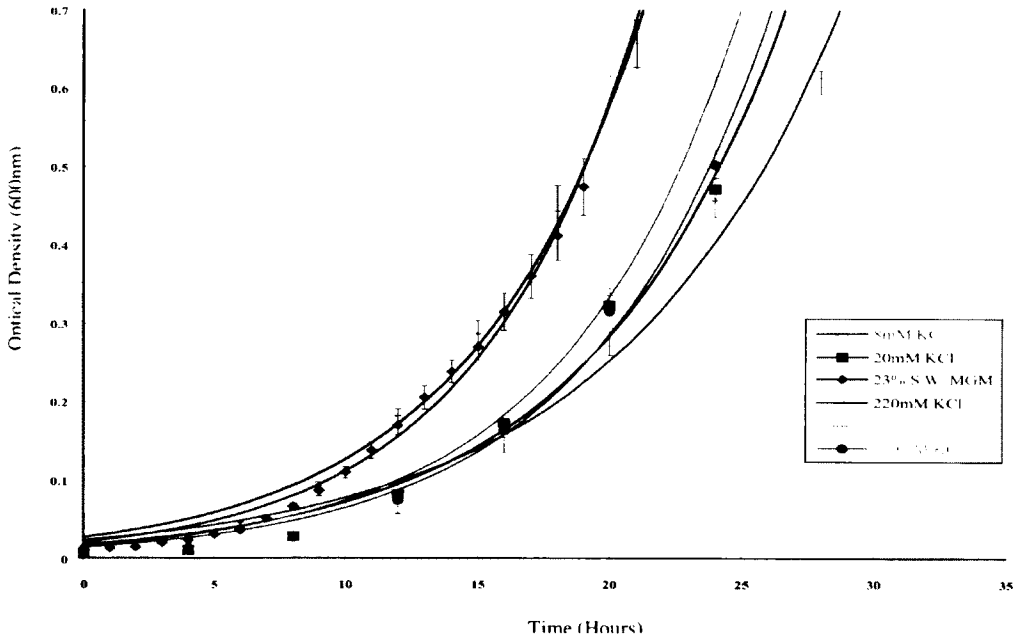

Figure 2.1. Exponential growth of *Haloarcula marismortui* under varying extracellular potassium concentrations. Curves were constructed by plotting OD_{600} against growth time. Optical density measurements were obtained once per generation time as determined by an experimental test curve for each condition. 23% S.W. MGM contains 100mM KCl. Error bars are representative of standard deviation obtained from a triplicate of biological triplicates. Curves were fit using the exponential growth equation. The resulting generation times can be found in Figure 2.2.

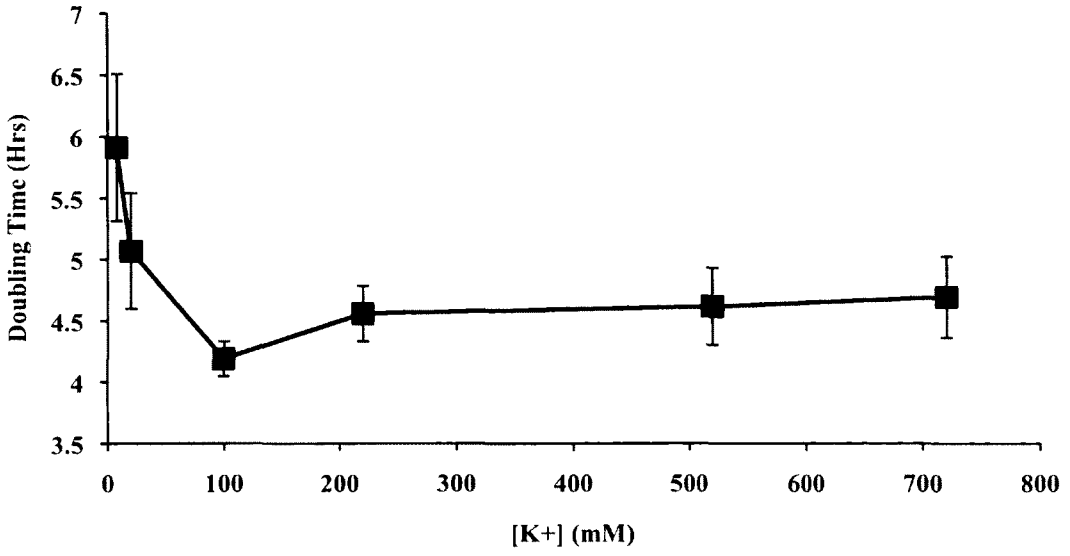

Figure 2.2 Average generation times obtained from the exponential growth curves (Figure 2.1) vs. extracellular potassium concentration. Error bars represent standard error obtained from the determination of cellular generation times via the exponential growth equation.

Using an adaptation of a previously reported method[112] as described above, the density of *Har. marismortui* cells was determined to be 1.20g/mL and confirms the result reported by Ginzburg *et al*[112]. Centrifugation of a layered suspension of cells resulting in 50% of cells (by visual estimate) passing through the high-density solution occurred in a solution comprised of 3.5M NaCl and 20% (w/v) sucrose which produced a scale density (determined by weighing varying volumes of the solution) of 1.20g/mL. Additionally, the construction of a standard curve (Figure 2.3) relating a volume-specific cell count to the optical density has allowed for an approximation of a cellular mass of $1.81\times10^{-12} \pm 0.144\times10^{-12}$g. When used with the density relationship (ρ=m/v), these two values have revealed a cellular volume of $1.509\times10^{-15} \pm 0.12\times10^{-15}$L (1.509 ± 0.12 fL) which allows for the determination of intracellular ion concentrations as described above.

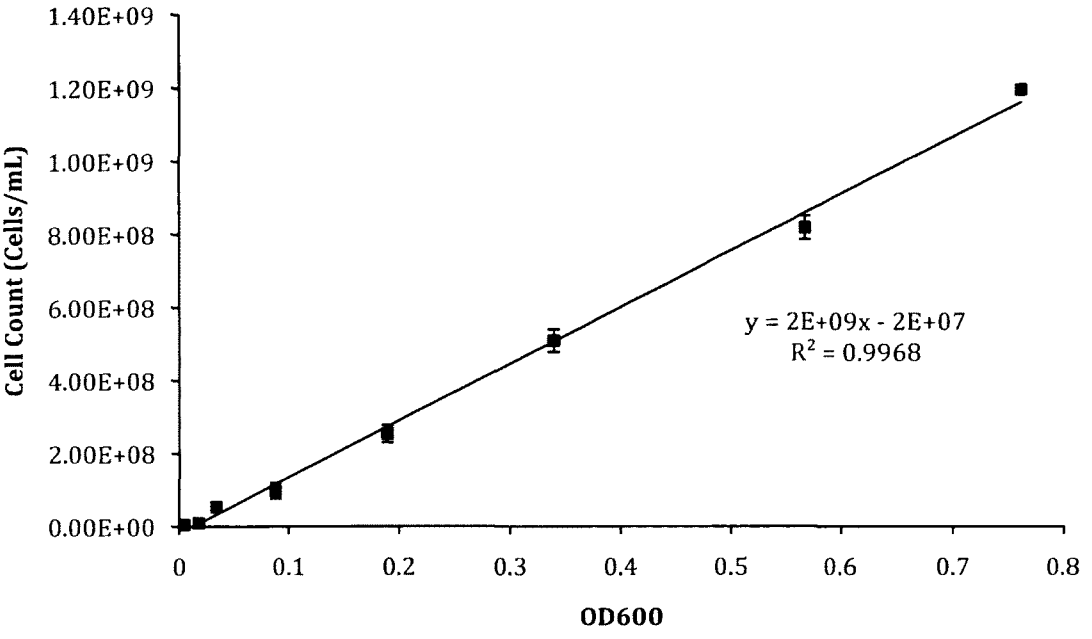

Figure 2.3. A standard curve relating cell culture density to optical density at 600nm. Cells were counted via standard haemocytometer as described in Section 2.2.4. A line of best fit was applied to the curve to obtain an equation (shown on graph) relating cell culture density in cells/mL to the measured OD_{600}. Error bars indicate standard deviation.

Intracellular ion concentrations were determined by trace metal analysis of cellular lysates using ICP-MS. Typically, when *Har. marismortui* is placed in water, cells lyse completely; however, when the cell pellets were lysed in 1% HNO_3, cellular materials were clearly visible in the lysates. To overcome this and ensure the 5mL lysates were suitable for ICP-MS analysis, lysates were sonicated prior to sample loading. Trace metal analysis was conducted to identify the concentrations of K^+, Li^+, Rb^+, and Cs^+ within the cell lysates. The ion concentrations within the 5mL lysates were then used in conjunction with the cell density standard curve and cellular volume described above to determine the intracellular concentrations for each of the latter ions under various growth conditions. Figure 2.4 illustrates variation in intracellular concentrations of K^+ with respect to changes in extracellular KCl concentration while Figure 2.5 and 2.6 illustrate variation of individual

and total ion concentrations, respectively, when cells are grown under 100mM LiCl, RbCl, or CsCl. Sodium ion was not included in these figures as the inability to wash cell pellets in an Na^+-free buffer lead to Na^+ concentrations that were excessive and highly unreliable (Section 2.4).

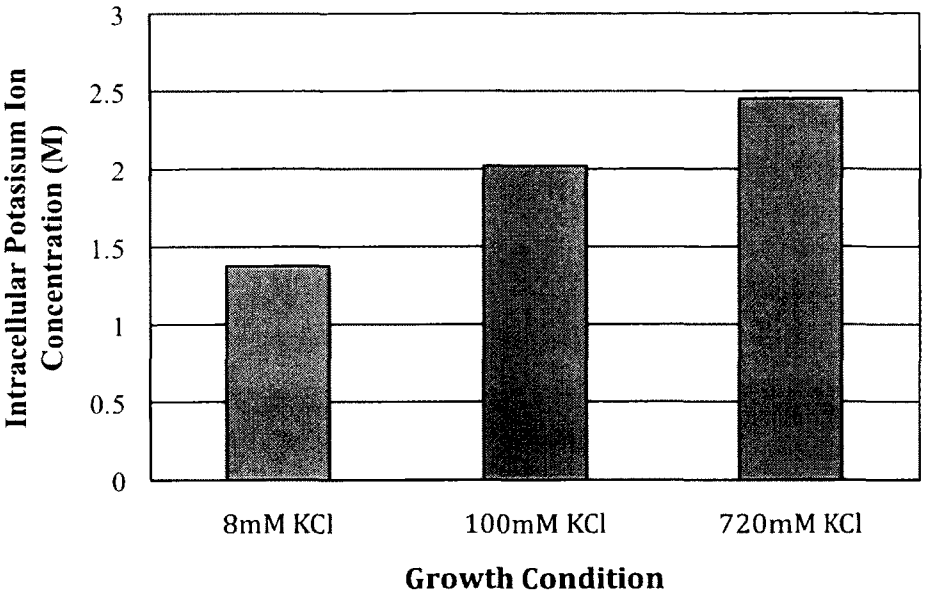

Figure 2.4. Intracellular concentrations of K^+ obtained via trace metal analysis with ICP-MS. Concentrations were obtained by lysing cells grown to balanced growth under 8, 100, or 720mM KCl in 5mL 1% HNO_3. The cell density standard curve (Figure 2.3) was used in conjunction with the optical density of the culture to determine the number of cells lysed. The number of moles of ion per cell were determined from ion concentrations obtained using ICP-MS and subsequently with the 1.509fL cellular volume described above to determine intracellular ion concentrations.

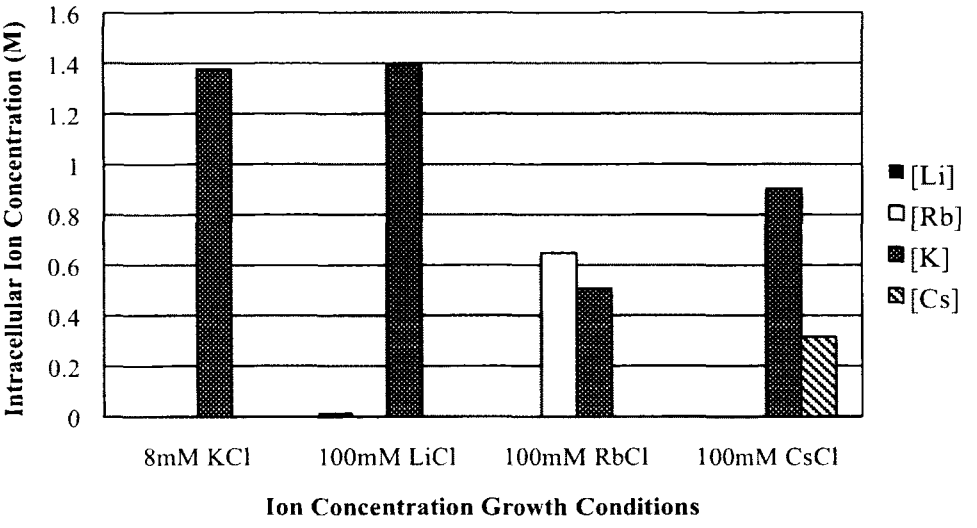

Figure 2.5. Intracellular concentrations of Li^+, Rb^+ and Cs^+ obtained via trace metal analysis with ICP-MS. Concentrations were obtained by lysing cells grown under 100mM LiCl, RbCl, or CsCl in 5mL 1% HNO_3. Alternative ion cultures were inoculated 1:100 with cells grown to balanced growth under standard conditions then incubated at 45°C until mid exponential growth was achieved. The cell density standard curve (Figure 2.3) was used in conjunction with the optical density of the cultures to determine the number of cells that were lysed. Concentrations obtained by ICP-MS were used to determine the number of moles of ion per cell, which in turn was used with the previously determined cellular volume (1.509fL) to determine intracellular ion concentrations. Intracellular ion concentrations observed with growth under 8mM KCl is provided for a comparison of concentrations typical of limiting potassium conditions.

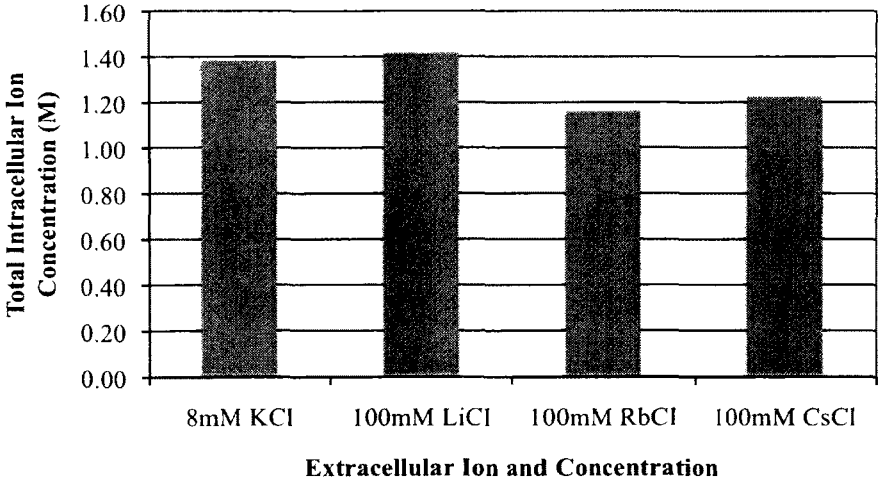

Figure 2.6. Total intracellular ion concentration of monovalent cations examined in *Har. marismortui* after growth on 100mM concentrations of the alternative monovalent ions of interest as described in Figure 2.5. Total ion concentrations are given as the sum of all individual ion concentrations. Intracellular ion concentrations observed with growth under 8mM KCl is provided for a comparison of concentrations typical of limiting potassium conditions.

2.4 Discussion

Haloarcula marismortui has exhibited an optimal generation time of 4.19 ± 0.14 hours when grown under standard, 100mM KCl, conditions at 45°C. As the KCl concentration in the media was decreased to the minimally attainable 8mM concentration, the cellular generation time increased to 5.91 ± 0.60 hours. The increase in generation time is likely indicative of an inability to sequester sufficient potassium to maintain proper intracellular functions or a result of changes in cellular energetics. Moving towards the opposite end of the spectrum, as extracellular KCl was increased to 220mM the generation time increased from optimal to 4.56 ± 0.23 hours and slowed only slightly to 4.69 ± 0.33 hours at KCl concentrations of 720mM. This suggests the demand for intracellular K^+ is likely being met and the slowed growth is potentially occurring due to ionic stress associated with the elevated total ion concentration and/or sub-optimal function of intracellular systems due to excessive intracellular K^+ concentrations. Both of the latter explanations for this observed increase in generation time are discussed in detail below. Previous cellular generation times have been determined for several members of the family *Halobacteriaceae* and have been reported as ranging from 1.5 to 3 hours [133] with some species exhibiting generation times in excess of 6 hours[137], thus placing the results described above on par with the existing literature.

Through a crude adaptation of the centrifugation technique described by Ginzburg *et al*[112], the density of *Har. marismortui* cells were determined. This process involved visually estimating the quantity of cells that were pulled through solutions comprised of varying concentrations of sucrose and sodium chloride during centrifugation in order to estimate the density. Since Ginzburg *et al*[112] completed this task with far greater precision, rough confirmation of their result was only required as the identity of the species they were

using was unknown. Confirmation of a cellular density of 1.20g/mL allowed for the determination of a cellular volume. Ginzburg *et al*[112] reported the volume of their isolate as $1.22mm^3$ (1.22μL) which is extremely large in comparison to the $1.509nm^3$ (1.509fL) volume determined by the methods described here. Considering human erythrocytes (red blood cells) have been reported as having a 90fL volume[140], the volume reported by Ginzburg *et al*[112] appears to be severely over estimated. Even though the attempt at eliminating the mass of the proteinaceous surface layer, proteins, and membranes was not done here, but was done by Ginzburg *et al*[112], our obtained volume appears to be a far more reasonable estimate considering *Har. marismortui* cells are substantially smaller than most eucaryal cells.

The intracellular ion concentrations of Na^+ that were observed were extremely elevated and would not intuitively be expected with the presence of a well-maintained Na^+ concentration gradient. Since the cell pellets were washed with the minimal potassium media prior to lyses in nitric acid, it is very possible a small excess of media was held within the pellet and on the pellet surface. It is extremely difficult to obtain an accurate measure of the intracellular Na^+ concentrations using the methods described here as cell pellets must be washed prior to trace metal analysis. All wash steps conducted here were completed using a high-sodium environment to prevent cell lysis due to abrupt changes in osmotic pressure, which is likely to skew results towards a higher observed concentration; therefore, the intracellular concentrations of Na^+ have been excluded from Figure 2.4. Although it was not used for this study, it may be possible to utilize a cell wash methodology that relies on polyethylene glycol (PEG) to maintain osmotic pressure. This would reduce the need for incorporating excess sodium chloride into the wash buffers and may produce slightly more reliable results.

Additionally, the variation we observed in the calculated intracellular Na^+ concentrations (data not shown) is very likely due to space-charge effects[141, 142] that can be observed when conducting ICP-MS analysis of lighter ions at high concentrations. These effects occur with the mutual repulsion of positively charged ions, which forces ions out of the argon plasma laterally within the instrument. This results in an overall lower sensitivity and occurs more noticeably at high ion concentrations. Space-charge effects are less apparent with ICP-MS analysis of heavier atoms due to their slower overall momentum within the argon plasma. The high concentration of NaCl in our standard growth media, and potentially within, or on the exterior of the cell pellets after washing, makes ICP-MS analysis of Na^+ exceedingly difficult and suggests the obtained results are highly unreliable.

The intracellular K^+ concentration within *Har. marismortui* cells have now been shown to be 2.02M under standard/optimal growth conditions (Figure 2.4). This concentration was reduced to 1.38M when grown under minimal potassium conditions (8mM KCl), which demonstrates this species ability to scavenge K^+ from its environment in order to maintain the suitable internal salt concentration that is required for osmoregulation. As suggested above, the increase in generation time for growth at this concentration of extracellular KCl could be indicative of an inability to sequester sufficient potassium to maintain the optimal function of intracellular systems. The decreased intracellular K^+ reported further supports this notion.

A change in available cellular energy production could provide a second possible explanation for the observed variation in cellular generation time as it is highly plausible that the cell is required to spend additional energy on maintaining a steady potassium gradient. Potassium efflux (leaking; spillage) has previously been shown to occur in a reversed manner through low-affinity potassium transporters such as the Trk symport

system in *E. coli* resulting in a futile cycling of potassium ions[143]. A similar occurrence in the homologous system in *Har. marismortui* would mean additional energy is required to maintain a gradient through the use of high-affinity, ATP-driven transporters under limiting extracellular K^+ concentrations since efflux would occur more readily under these conditions. The increased energy requirement to maintain a steep K^+ gradient would leave less energy available for other cellular process thus cellular generation time would be expected to increase. This is supported by the dramatic increase in generation time that was observed under limiting extracellular K^+ concentrations and only a mild increase in generation time under extremely elevated extracellular K^+ concentrations.

The Trk system does not hydrolyze ATP to function, but rather utilizes it for regulatory purposes only (reviewed by Oren *et al*[136]). As pointed out by Oren *et al*[136], it is possible that an increased demand for K^+ uptake through the Trk system could immobilize a substantial quantity of ATP if regulation occurs through the binding of ATP in a transporter active site or allosteric site. This in turn would limit the concentration of ATP available for other cellular processes. Oren *et al*[136] also state that, at the time of their review on the bioenergetic aspects of halophilism, a high-affinity potassium transport system similar to the Kdp system found in *E. coli*[136, 143] had not yet been identified in any of the haloarchaeal species. A survey of available literature suggests that although a haloarchaeal high-affinity system such as this remains elusive, its non-existence remains to be proven. Therefore, a high-affinity potassium transport system or significant immobilization of ATP by the Trk system could still potentially be the cause of reduced energy stores available for normal intracellular processes which would explain the severely increased generation time observed under limiting K^+ concentrations.

Alternatively, the intracellular K^+ concentration was observed to increase to 2.45M when grown in the presence of 720mM KCl. One possibility for this increase could be that *Har. marismortui* is unable to regulate the internal salt concentration in response to elevated extracellular KCl concentrations and will sequester additional ions if they become available, thus producing elevated, sub-optimal internal salt concentrations. Because the latter energy requirement would not be expected to exist when cells are growing under high K^+ concentrations, the argument for a change in cellular energetics due to potassium efflux is far less plausible in this case. The gradient under these high extracellular K^+ concentrations is only a fraction of the strength of the gradient formed under standard conditions and a far smaller fraction of the strength of the gradient formed under limiting potassium conditions. It is therefore far more probable that the more subtle increase in generation time that was observed can be attributed to sub-optimal function of intracellular systems due to an inability to regulate K^+ uptake as described above. Regardless of which of these suggestions is actually occurring, an additional stress, whether intracellular, extracellular, energetic, or a combination thereof, is being exerted on the cells under these conditions resulting in an increased cellular generation time.

It is also possible observed increases in generation time under elevated extracellular K^+ concentrations is in response to the increased ionic stress being exerted on the cells in this media. Since extracellular NaCl was not decreased when KCl was increased, a constant total ion concentration was not maintained. Countless publications have described the numerous signalling pathways that are affected by ionic stress in both bacterial and eucaryal cells indicating ionic stress such as that placed on cells in this study may have a profound effect on cellular generation time. Further investigation of ionic stress responses in *Har. marismortui* is still required.

Rubidium has been previously used as an aid to ion transport studies after another halophilic archaeon, *Haloferax volcanii*, was shown to uptake Rb^+ during K^+ starvation. This attribute has been exploited in an attempt to estimate the quantity of permanently bound K^+ within these cells by forcing a maximum quantity of mobile K^+ ions out of the cell via exchange with Rb^+ [137]. The publication in which this was reported suggests that 50% of the K^+ in *Hfx. volcanii* is exchangeable with Rb^+ indicating up to 1.8M concentrations of K^+ is free and non-bound while the remaining K^+ (up to 1.5M) is tightly bound within the cells[137]. The findings outlined above appear to be agreeable with this statement and appear to confirm this finding in a second halophilic species. When grown in media containing 100mM RbCl the intracellular K^+ concentration dropped to 0.51M while intracellular Rb^+ increased to 0.65M. Though this would suggest confirmation of the previous report, the concentration of contaminating K^+ in the NaCl used for media preparation must also be considered before these findings can be confirmed. Even with the use of NaCl that has been assessed as 99.999% pure by trace metal analysis, potassium remains present in the stock sodium salts at concentrations 5mg/Kg NaCl and to unknown concentrations in the majority of other reagents used. This corresponds to a conservative estimate of an excess of 8mM potassium in the growth media. As this is the lowest attainable K^+ concentration, the possibility that the apparent exchange of Rb^+ with mobile K^+ ions, as observed in Figure 2.5, is actually an accumulation of both K^+ and Rb^+ from the media can not be ruled out. The previous suggestion that the remaining intracellular K^+ concentration is non-mobile and bound within the cell (as per observations in *Hfx. volcanii)*, though supported by the neutron scattering experiments described in section 2.1.2.1 above, can not be confirmed until a growth media can be prepared with substantially lower K^+ concentrations that will allow for improved evaluation of ion exchange. Although

these studies were not conducted in *Har. marismortui*, it does not appear likely that the authors of that publication accounted for the presence of contaminating potassium in their salt stocks. It is instead more likely that *Hfx. volcanii* and *Har. marismortui* are accumulating both K^+ and Rb^+ without, or with minimal bias as a means of osmoregulation.

A similar result is observed when cells are grown in the presence of caesium, albeit to a lesser extent. Figure 2.5 shows Cs^+ is sequestered internally to concentrations of 0.32M and the intracellular K^+ concentration falls to 0.90M when *Har. marismortui* is grown in the presence of 100mM CsCl. This is again quite likely an unbiased accumulation of both ions as a means of maintaining osmoregulation. The similar chemistries, and potentially similar hydration of these ions may allow for transport of Cs^+ into the cell via established K^+ transporters. The ATP-regulated K^+ transporter that is similar to the Trk system found in *Escherichia coli*[136, 137] may provide additional merit to the latter statement. Previous work has shown that the Trk system, specifically the low-affinity K^+ uptake protein *trkd*, is capable of transporting Rb^+ and Cs^+ in addition to K^+ in some bacterial species[144, 145], but only K^+ or Rb^+ in others. The presence of this system in members of the family *Halobacteriaceae* makes the transport of the alternative ions via this transporter quite plausible. The high chemical and size similarity of Rb^+ to K^+ (as observed from the well established trends in the periodic table) explains the near identical intracellular concentrations observed of these two ions when grown in 100mM RbCl. The observed intracellular concentration of Rb^+ is slightly higher than that of K^+ which is likely due to a larger abundance of the ion. Alternatively, the lower concentration of Cs^+ in comparison to K^+ observed when cells were grown in 100mM CsCl can be explained by the substantially larger size of the Cs^+ ion making transport of this ion more difficult. Similarities in the K^+, Rb^+, and Cs^+ hydration shells could also explain the observed patterns in the uptake of

these ions. Previous reports have shown that specific hydration patterns are required within the transport channel in order for potassium transport channels to accurately select K^+ ions as they are moved across the cell membrane in *E. coli*[146]. Therefore, even though the size of the Rb^+ and Cs^+ ions differ from that of the K^+ ion (Cs^+ to a far greater extent) which should be a preliminary selective bias, the potential for these ions to form the proper hydration shell after the initial stripping of water from the ions at the transporter entrance could be the cause for reduced selectivity.

The accumulation of Li^+ as an alternative to K^+ does not appear to occur when grown in 100mM LiCl; however, Li^+ concentrations observed in Figure 2.5 (0.02M) may not be overly accurate due to the space-charge effect that occurs with lighter atoms during ICP analysis (described above). This may suggest the actual intracellular concentration could possibly be substantially higher. That being said, because a decrease in K^+ concentration is not observed in these cells in comparison to those solely grown under 8mM KCl , as was observed with growth in the presence of RbCl and CsCl (See Figure 2.5), it is unlikely *Har. marismortui* is sequestering Li^+ under limiting potassium conditions. Given the available data, it seems more plausible that cells are scavenging K^+ from the extracellular media when grown in the presence of Li^+ and the ion is excluded from the cell exclusively presumably due to the substantially smaller size of the ion.

2.5 Conclusion

The resilience and adaptability of *Haloarcula marismortui* to respond to changes in its native environment has been clearly demonstrated through the evaluation of growth characteristics during potassium stress. Cellular generation times increase substantially as extracellular potassium reaches minimal concentrations and these changes in generation

time are reflected by changes in intracellular ion concentrations. Under minimal K^+ growth conditions, intracellular K^+ is reduced by approximately 30%, which is presumably near the lower concentration limit required to maintain internal cellular functions. The drastically increased generation times observed under limiting K^+ concentrations has been attributed to a decrease in energy available to maintain cellular processes due to the amount of additional ATP that is either being hydrolyzed by a high-affinity potassium transport system or immobilized by the Trk system during regulation of that transporter.

Alternatively, intracellular K^+ concentrations increase by nearly 25% when cells are grown on 720mM KCl. Cellular generation times remain steady, but sub-optimal, as K^+ concentrations increase toward total salt saturation. We have suggested that the mild increase in observed generation time under these conditions is likely due to poor regulation of K^+ uptake. This is likely to result in sub-optimal operation of cellular systems as intracellular K^+ concentrations begin to affect protein folding and function, but not to a point of complete detriment. Although total ionic stress could be the result of the observed increase in generation time, this suggesting needs to be further evaluated.

On a final note, since the dry mass of the cell pellets was not measured in order to account for the mass of cellular components, the calculated cellular volume is likely slightly overestimated. Therefore, the intracellular concentrations reported above are likely to better serve as a report of the total concentration of each ion bound by the cell as a whole, and should not be regarded as only the concentration of free ions in the cell water.

Chapter Three

Identification of Novel RT-qPCR Reference Genes in *Haloarcula marismortui*

3.1 Introduction

3.1.1 A Brief History of RT-qPCR

Reverse-transcription quantitative Polymerase Chain Reaction (RT-qPCR) is becoming an increasingly popular analytical method for the analysis of gene expression. As the name suggests, the technique is a modification on the traditional Polymerase Chain Reaction (PCR). In 1992, Higuchi *et al*[147] modified the PCR in a manner that allowed for the detection of specific double-stranded DNA amplicons without opening the tube. This was accomplished by conducting the PCR in the presence of ethidium bromide and monitoring the increase in fluorescence via fibre optic cable and fluorometer as the reaction progressed. Higuchi *et al*[148] improved this method in 1993 by utilizing a video camera to continuously monitor the fluorescence increase throughout the PCR. The kinetics of ethidium bromide fluorescence accumulation were found to be directly related to the number of DNA template copies in the reaction mixture, thus providing a basis for modern real-time monitoring of the PCR.

In 1996, Gibson *et al*[149] utilized reverse-transcription and a target-specific, dual-labelled fluorogenic probe containing a reporter dye and a quenching dye described earlier by Livak *et al*[150] in a real-time mRNA quantification assay. When the probe was intact the reporter dye would be quenched due to the close proximity of the quenching dye; however, upon probe hydrolysis via DNA polymerase 5'-exonuclease activity[151], the quencher would be released from the reporter and fluorescence would be observed. This allowed for real-time monitoring of a specific cDNA target sequence during PCR amplification and the

subsequent quantification of a specific mRNA within a sample[149]. Modern RT-qPCR probes, chemistries, methodologies, instruments, and data analysis has since been extensively reviewed by S.A. Bustin[152].

In 1999, Gygi *et al*[153] determined that the number of mRNA transcripts in *Saccharomyces cerevisiae* as obtained by RT-qPCR did not correlate to expressed protein levels. This showed that mRNA levels are not always directly related to the expression of their translated protein counterpart and thus, one of the major pitfalls of RT-qPCR data interpretation was identified. Bustin *et al*[154] has extensively reviewed additional pitfalls pertaining to various RT-qPCR methodologies and trends that have been observed in more recent years with respect to the reporting of RT-qPCR results. Of particular importance is the linearity of the reverse-transcription step. As explained by Bustin *et al*[154], the synthesis of cDNA via reverse-transcription can be conducted using random, deoxythymidine oligonucleotide (oligo-dT), or target-specific primers, each of which produce differing yields, variety, and specificity. The calculated mRNA copy number obtained by RT-qPCR can vary widely based on the chosen methodology. Many of the remaining major pitfalls associated with this highly sensitive technique stem from the analysis of obtained data and how it is reported. Bustin *et al*[132] attempted to clarify this as described below.

3.1.2 The MIQE Guidelines

Since the introduction of RT-qPCR investigators have been utilizing a diverse range of reagents, protocols, data analysis methods, and publication formats. As described above, there are pitfalls associated with these methodologies due to typically unacknowledged variables throughout the RT-qPCR process[154]. In 2009 Bustin *et al*[132] observed the enormous lack of consistency across publications using qPCR data and specifically

identified several recurring technical deficiencies limiting reproducibility or assay performance. These deficiencies include inadequate sample collection, preparation, quality, and storage, poor choice and optimization of primers and probes for the PCR, and the generation of potentially misleading results due to inappropriate analysis of data. In an attempt to bring consistency to future publications utilizing RT-qPCR, Bustin *et al*[132] proposed a set of guidelines for the Minimum Information for publication of Quantitative real-time PCR Experiments (MIQE Guidelines). This document clarifies terminology as it applies to the technique and the proper use of that terminology, as well as proposing an "industry standard" for methods of primer design, sample preparation, RNA integrity analysis, controls, reverse-transcription, qPCR assays, data analysis, and publication of RT-qPCR results. If followed, all published data would follow a consistent, standard minimum format constructed from reliable and reproducible data.

The MIQE guidelines outline several specific considerations that must be taken into account while designing, conducting, and publishing the results of RT-qPCR assays. These experimental and publication considerations, starting with sample preparation and ending with the analysis of data (Section 3.1.3.1 – 3.1.2.7), are briefly outlined below as originally proposed by Bustin *et al*[132]:

3.1.2.1 Sample Preparation

The detailed reporting of tissue sample collection and processing is a necessity. Publications should clearly state where tissue samples were obtained and whether or not they were immediately processed. If samples are stored for any length of time it is necessary to report the method, length, and conditions of preservation. The extraction of nucleic acid is a critical step as extraction efficiency depends on a number of factors

including the amount of biomass initially processed, physiological status, and adequate sample homogenization. The details of the method of extraction used should also be clearly reported. The extent of genomic DNA contamination should be tested and reported along with the threshold cut-off criteria that was used to determine the amount of tolerable contamination. Reporting of the type of DNase and the corresponding reaction conditions used is essential if the RNA sample was treated with such.

3.1.2.2 RNA Quantification and Quality Assessment

Quantification of RNA in extracted samples is critical when conducting RT-qPCR assays and it is necessary to report the methods used to measure the concentrations and RNA quality in detail. Most common methods of RNA quantification will produce varying results when compared to alternative methods making it difficult to compare data obtained by one method to those obtained from another. With several methods available such as spectrophotometery, micro fluidic analysis, or capillary gel electrophoresis, it is recommended that only a single method be used to maintain consistency in results. The preferred method utilizes fluorescent RNA-binding dyes which are ideal for the detection of low RNA concentrations. The quality assessment of RNA templates is also an essential step and can only be bypassed when the quantity of extracted RNA is not sufficient enough to allow for quality assessment or when extraction and RT-qPCR steps are being performed as a continuous, single-tube experiment. Gel electrophoresis evidence of RNA integrity should be reported at minimum.

3.1.2.3 Design of Primers and Probes

Several *in silico* tools, such as BLAST, are useful aids in the design of primers for RT-qPCR assays. Publications must provide primer sequence information as well as an assessment of primer specificity. The structure of the target nucleic acid should also be considered. Secondary structures may have an impact on reverse-transcription and PCR efficiency. *In silico* nucleic acid-folding tools such as mFold should be utilized when considering the positioning of primers or probes on the target molecule. This folding structure data should also be made available to reviewers at the time of manuscript submission. The primer supplier's lot information and experimental validation criteria are also required to be provided in publications. Predicted primer homology to unexpected targets or pseudogenes should be provided to reviewers as aligned sequences; however, direct experimentation must be conducted to validate primer specificity during assay optimization. Additionally, although *in silico* tools for the prediction of primer annealing temperatures (T_a) exist, the optimum temperature for annealing should still obtained experimentally. Although primer optimization is not commonly practiced, it is critical as poor optimization can have substantial effects on assay quality.

3.1.2.4 Reverse Transcription: Synthesis of cDNA

Reverse transcription can introduce substantial variation in an RT-qPCR assay. A detailed description of the reagents and protocol used in the formation of cDNA from RNA must be provided. This description should include the quantity of RNA that was reverse-transcribed, the priming strategy, the type of enzyme used, volume and temperature of the reaction, and duration of the reverse-transcription step. This step should be carried out in

duplicate or triplicate and the total RNA concentrations should be identical across replicates.

3.1.2.5 The qPCR Assay

The detailed reporting of RT-qPCR assay information is required in publications. This information must include database accession numbers for all target and reference genes, the locations of each primer and probe on the exon, sequences and concentrations of each oligonucleotide, and the identity, position, and linkages of any dyes or modified bases used. Reaction conditions including the identity and concentration of the polymerase, the exact compositions of buffers, the Mg^{2+} concentration, and the reaction volume are also required in publications. In addition to the above investigators must identify the instrument being used, document the cycling conditions, and report the identities and manufacturers of single tubes, strips, or plates as these consumable items can affect thermal cycling. The method of sealing used should also be reported when using plates.

3.1.2.6 Controls

As with any scientific experimentation, the use of controls in RT-qPCR is required. These should include no-reverse transcription controls during the reverse-transcription step as well as the use of NTC's should be included on each plate or with each sample set when running RT-qPCR assays. Conditions for data rejection should be established with these controls in mind. Positive controls will allow for monitoring of variation between assays and must be used when calibration curves are not being conducted for each run. Investigators must be aware that most quantitative RNA data are relative rather than absolute. Standardization of results is required and can be achieved through the use of

reference genes. The assessment of the validity of experimental results must consider the relative quantification reference and whether or not it is appropriate.

3.1.2.7 Analysis of Data

In order for RT-qPCR data to be considered viable, assay performance characteristics must be assessed. These characteristics include PCR efficiency, determination of the linear dynamic range of the RT-qPCR assay via the construction of standard curves, and determination of the limit of detection. Once performance characteristics are established it is important to recognize appropriate Threshold Cycle values (C_T) and normalize results in an appropriate manner. Helpful information on all of the latter can be found in the reviews by Bustin [152] and Bustin *et al* [154] mentioned above.

3.1.3 Considerations for RT-qPCR Studies in Archaeal Systems

3.1.3.1 Archaeal Introns and Intron Processing

In order to conduct RT-qPCR studies the characteristics of archaeal RNAs must be considered. The *Archaea* lack nuclei and their genome and gene structure are most like that of the *Bacteria*; however, their DNA replication, transcription, and translation mechanisms are more similar to the *Eucarya* (reviewed by Klug *et al*[155]). As in the other two domains, introns have been found within exon-coding regions of archaeal genes. Introns are transcribed with the pre-mRNA transcript by an RNA polymerase most similar to the RNA polymerase II and III found in the eucaryotic domain of life (reviewed by Langer *et al*[51] and Hirata *et al*[53]) before being processed out of the mature transcript by endonuclease activity[61]. Though Group I and Group III introns are yet to be detected, Group II introns have been identified only in the *Methanosarcinaceae* which fall into the kingdom

Euryarchaeota[156]. Several rRNA and tRNA genes across the archaeal domain, as well as a single protein-coding gene in a small handful of archaeal species[60], have been found to contain small introns of 14 to 106 nucleotides which appear to be related to the eucaryotic nuclear tRNA introns[155].

As in the other domains, archaeal introns are removed by endonuclease excision after transcription[155]. Many archaeal species have introns in their tRNA and rRNA genes, all of which form a conserved secondary structure comprised of a small bulge loop followed by a short helix and a second bulge loop on the opposite side of the base paired intron sequence. Endonuclease excision occurs at a site specific location within either bulge loop of these bulge-helix-bulge (BHB) motifs and mutations in which these locations are removed or made inaccessible prevent splicing from occurring entirely[61]. The addition of nucleotides to the helix between the bulge loops causing them to be further separated from one another results in reduced cleavage efficiency; however, cleavage accuracy remains intact[61]. Slippage in the cleavage sites will occur, however, when an additional nucleotide is added to one of the bulge loops though this is a minor occurrence and cleavage accuracy remains in the majority of cleavage products[61]. It is believed the archaeal intron-processing endonuclease requires a structural recognition pattern that allows for proper alignment of the intron with the enzymes active site and, though cleavage site measurement appears to play a role in cleavage efficiency, nucleotide recognition may be utilized to identify the intron splice sites[61]. BHB motifs have now been identified at a number of locations within tRNA genes from a wide array of archaeal species. The majority of these introns conform to the BHB motif described above while a small minority exist as relaxed motifs that can be cleaved by the tRNA processing endonuclease found in *Crenarchaeota* which is thought to be a more primitive variant[59].

In addition to the well established tRNA introns, Watanabe *et al*[60] reported the presence of small introns in the archaeal protein-coding gene homologue, CBF5 (Centromere-binding factor 5), of four archaeal species in 2002. Computational analysis predicted the introns would form a BHB secondary structure motif similar to those found in the archaeal rRNA and tRNA introns with splice sites predicted to be located in either bulge. The ligation of the splice sites was confirmed by sequence analysis of reverse-transcribed cDNA confirming the presence of the first known intron in an archaeal protein-coding gene[60].

3.1.3.2 mRNA Characteristics in Archaeal Systems

Messenger-RNAs found in *Bacteria* are typically polycistronic and contain a 5′-triphosphate group[155] that has been shown to protect mRNA from the endoribonuclease, RNase E, thus prolonging mRNA decay[157]. Bacterial mRNAs have also been shown to contain short polyadenylated (poly(A)) tails on the 3′-end[155] which play a role in mRNA degradation signalling[158]. On the other side of the spectrum, the *Eucarya* possess mRNAs that are typically monocistronic and contain 5′-methylated guanosine caps with long, non-templated poly(A) tails that provide stability to the message[155]. Archaeal mRNAs are often found as polycistronic transcripts[155] and several species have been identified as containing poly(A) tails which have been observed to reduce the stability of the transcript[159, 160]. No 5′ modifications have been identified on archaeal mRNAs[159] suggesting archaeal messages are more similar to those found in *Bacteria* over those found in *Eucarya*. Currently the mechanism of mRNA decay in the *Archaea* is poorly understood and several studies

investigating archaeal mRNA stability have provided results showing great variability in mRNA half-life and stability across a range of species [155].

3.1.4 Previous Studies Utilizing RT-qPCR in Archaeal Systems

RT-qPCR is now becoming a more widely utilized method for the analysis and quantification of archaeal RNAs. The technique has been used in investigations ranging from the determination of ammonia-oxidizing archaeal species abundance in soils[161] to identifying changes in expression between the multiple rRNA operons of the halophilic archaeon, *Haloarcula marismortui*[162]. Since archaeal mRNAs do not possess long poly(A) tails like their eucaryotic counterparts, and the poly(A) tails available are involved in mRNA decay signalling, the use of poly(T) primers for reverse-transcription is not appropriate. Instead, reverse-transcription has previously been conducted using random hexameric primers[162] or primers specific for a target gene[163] with apparent success. Reverse-transcription has also been conducted in archaeal halophiles as an initial step to RNA Arbitrarily Primed PCR (RAP-PCR)[164]. This method of reverse-transcription, which is highly similar to the use of random hexameric primers, utilized arbitrary 10-mer primers with G+C contents of 70% and 80%[164] which may assist in primer annealing as the elevated G+C content observed in many archaeal species is accommodated for. First strand cDNA synthesis was then performed[164]. Though this method was not used for downstream RT-qPCR applications it provides an additional alternative that may assist in achieving improved reverse-transcription results.

Quantitative reverse-transcriptase PCR studies have been conducted by Lopez-Lopez *et al*[162] to determine differential expression between *Har. marismortui*'s three rRNA operons. The organism possesses two rRNA operons that are highly similar and a third that

contains a substantial number of single nucleotide polymorphisms. Analysis by RT-qPCR revealed that the unique operon was over-expressed as temperature increased suggesting the divergent operon is expressed as an adaptive measure. Lopez-Lopez *et al*[162] were able to successfully use random hexameric primers for reverse-transcription. After determining the *polA1* gene (encoding the small DNA polymerase II sub-unit), which had been used in other archaeal RT-qPCR analysis, had variable expression in *Har. marismortui*, the ratio of expression between the two similar rRNA operons under a standard laboratory growth condition was utilized as a reference to normalize the expression of the unique operon. This was done based on an argument that rRNAs are the most commonly used reference genes in procaryotes and since the two operons were identical it should be possible to normalize against the expression ratio between the two operons if the ratio is taken to be 1[162]. To the best of our knowledge, this is the only study that has conducted a RT-qPCR assay to assess expression of any kind in *Har. marismortui*. Although it is ground-breaking, this investigation does not meet the reference gene requirement as outlined by the MIQE guidelines. Since *Har. marismortui* is a widely used model organism for the eucaryotic ribosome, the proper identification of a useable subset of reference genes for RT-qPCR assays is going to be invaluable and would benefit the current research being conducted with respect to ion transport in *Har. marismortui* described above. Moreover, there does not appear to be any reason as to why the minimum standards required by the MIQE guidelines [132] can not be followed up to, and including, the confirmation of reference gene stability and the continuous quality control of samples.

3.1.5 Study Specific Objectives

Publications released in the last 2 years reporting RT-qPCR assays conducted during investigations pertaining to archaeal species can still be found that do not yet appear to fully conform to the MIQE guidelines since their conception in 2009[165, 166]. If researchers are in fact conducting assays with these guidelines in mind it is not always apparent in published materials as substantial quantities of information deemed as required for publication by Bustin *et al* [132] is not always clearly stated. The most obvious missing information is a lack of reported reference genes or alternative method for standardizing RT-qPCR data. In the event reference genes are being used and their use is being reported, the likelihood of data being normalized against more than one reference gene remains minimal. Although this lack of reference genes is not the only concern with regards to published RT-qPCR assay data it is one of the major concerns. In order to properly address this matter adequate reference genes need to be identified in a variety of archaeal species.

3.1.5.1 Specific Considerations for RT-qPCR Studies in *Haloarcula marismortui*

The genome of *Har. marismortui* is comprised of nine circular replicons totalling 4.28Mb. These replicons are organized into one large 3.13Mb chromosome, a substantially smaller 288Kb chromosome, and 7 plasmids ranging from 33Kb to 410Kb in size[167]. As with most other archaeal species, *Har. marismortui* contains short introns within its tRNA genes which are spliced out of the mature mRNA transcript via tRNA endonuclease cleavage at the BHB motif[168]. Moreover, the halophilic archaea are unique from most other archaeal species in that the 3′-ends of their mRNA lack poly(A) tails entirely[169] which leaves the mechanism of halophilic mRNA decay a mystery.

As mentioned above, the lack of poly(A) tails on *Har. marismortui*'s mRNA leaves random hexamers or gene specific primers as the only employable option for priming cDNA synthesis via reverse-transcription. Due to the number of genes being investigated, random hexamers are used as this methodology is more cost effective. Additionally, the placement of probes, if being used, across introns boundaries is not necessary as introns do not appear in protein coding genes as mentioned above.

3.2 Methods

3.2.1 Identification of Candidate Reference Genes and Design of RT-qPCR Primers

Candidate reference genes (Table 3.1) were selected on the basis of their use in other organisms or the likelihood of a uniform expression pattern with respect to their use in metabolic pathways. Candidates that had not been previously established in other species were selected based on their location within their respective metabolic pathway and on the likelihood of being uniformly expressed due to the number of 'inputs' feeding said pathway.

Table 3.1. Candidate reference genes selected on the basis of prior use in RT-qPCR experimentation or probability of uniform expression expected due to location of gene product involvement in metabolic pathways.

Gene	Loci	Gene Product
-	16S rRNA*	16S rRNA*
gapB	rrnAC2262	GAPDH
rpoA	rrnAC2428	RNA polymerase alpha sub-unit
rpoB	rrnAC2430	RNA polymerase beta sub-unit
polA2	rrnAC2691	DNA polymerase II large sub-unit
pykA	rrnAC0546	Pyruvate Kinase

*The *Har. marismortui* genome contains three rRNA operons, each of which has a single 16S rRNA. Expression of this RNA will be measured as a combined expression of 16S rRNA from operons rrnA and rrnB as primer pairs can not be made to accurately distinguish between these two regions. The third 16S rRNA on operon rrnC has been previously identified as having differential expression[162].

All primers were designed using the Beacon Designer 7 software package (Premier BioSoft). The locations of candidate reference gene sequences were identified using the UCSC Archaeal Genome Browser[85, 86, 170, 171]. Location information was then used in conjunction with the NCBI *Har. marismortui* chromosome I whole genome sequence to obtain each specific gene sequence. The BLAST feature[170] within the software suite was used to search each nucleotide sequences against the *Har. marismortui* genome in order to maximize primer specificity and prevent annealing to homologous sequences within the genome. The software's secondary structure search option was used to determine regions of secondary structure across each gene. This cross homology and structural information was used to design primer pairs (Table 3.2) that produce amplicons ranging from 150-250 base pairs in length (G+C = 45-65%) using the Beacon Designer 'avoid cross homology' and 'avoid secondary structure' options. Longer amplicons were only generated when primer design parameters for the desired lengths could no longer be relaxed further regarding nucleotide runs, G+C content, annealing temperature, and hairpin or primer dimer binding energies.

Table 3.2. Primer pairs for qPCR assays of subset #1 candidate reference genes. All primers were designed using the Beacon Designer 7 software package (Premier BioSoft) and supplied by Integrated DNA Technologies.

Gene	*Haloarcula* Loci	Sense Primer	Anti-Sense Primer	Amplicon Length (bp)
16S rRNA	Operon rrnA/B	5'- GGTTGACGACTTTACTCG	5'- GTCATAGCCATTGTAGCC	242
*gapB**	rrnAC2262	5'- AACTACGAGAAGGCAGTC	5'- CAACTGTGATGGATTCGG	493
rpoA	rrnAC2428	5'- GAACTCCAGATTCAGGTC	5'- CACACCAATCCGAAGTTC	165
rpoB	rrnAC2430	5'- ATCTCTGCGAGTTCCTGT	5'- GTTGACCGGACAGACTTC	239
polA2	rrnAC2691	5'- GAACTGGAAGGACAGATG	5'- GCGTAGTAGACGTTGATG	230
pykA	rrnAC0546	5'- GATGCTTGACTCGATGGT	5'- ATTTCGCCGATTTCAGGG	338

*The *gapB* gene was excluded from the study due to the excessive size of the amplicon produced by the qPCR primers. These were the only primers found that fell within allowable primer parameters using the Beacon Designer 7 software.

3.2.2 Preparation of *Haloarcula marismortui* Cell Cultures

Har. marismortui cells were grown in 23% S.W. MGM[138, 139] as described in chapter two. Cells were continuously sub-cultured at mid-exponential growth (OD_{600} = 0.4-0.6) as a means of maintaining continuously generation cultures as before. Cells for reference gene assessment were obtained by sub-culturing cells in balanced growth to media maintained under varying temperature and potassium ion concentrations which included growth in 23% S.W. MGM at 37°C, 45°C, and 55°C and in media containing 20mM KCl and 720mM KCl at standard temperature (45°C). Three biological replicates under each condition were maintained. Cells (OD = 0.4-0.6) were pelleted by centrifugation once balanced growth under each condition had been obtained and were used immediately for RNA extraction.

3.2.3 Extraction and Assessment of RNA and Subsequent cDNA Synthesis

As a means of minimizing freeze-thaw cycling of RNA extracts and to preserve RNA integrity, each of the following procedures were completed sequentially and immediately after one another. RNA was extracted from each biological replicate (above) using the Qiagen RNeasy kit. Cells in mid-exponential, balanced growth were centrifuged at 13,000rpm (1.5mL/2 minute sequential spin; total vol: 4.5mL) and media was removed completely by pipette. The cell pellet was resuspended and lysed using the RNeasy kit's guanidine thiocyanate containing buffer, RTL, then homogenized using a QiaShredder spin column (Qiagen) as per manufacturer's instructions. Ethanol (70% v/v) was added to the homogenization and ran over the RNeasy spin column as per the manufacturer's protocol then eluted with nuclease-free water. A preliminary assessment of collected RNA purity was conducted post purification via agarose gel electrophoresis. RNA samples were

denatured by incubation in the presence of a formaldehyde/formamide cocktail mix containing ethidium bromide (40mM MOPS, 1mM EDTA, 1mM sodium acetate, 50% (v/v) formamide, 6.5% (v/v) formaldehyde, 0.1ug/mL ethidium bromide) at 75°C for 15 minutes. Denatured samples were run on a 1% (w/v) agarose-Tris/Acetate/EDTA (TAE) gel then visualized over UV light (Figure 3.1). Following preliminary assessment, RNA samples were digested using the Ambion Turbo DNA-free™ kit as per the manufacturer's instructions. Purified RNA samples were stored at -80°C for up to 7 days before quantifying and assessing integrity. RNA was quantified via the Qubit® fluorometric system (Invitrogen) and integrity was assessed using the Experion Micro-Capillary Electrophoresis system (BioRad). Both methodologies are preferential as per the MIQE Guidelines[132].

Single strand cDNA synthesis was conducted using random hexameric primers as per Lopez-Lopez *et al*[162]. cDNA synthesis was completed in 20µL reactions containing 500ng RNA using Moloney Murine Leukemia Virus (M-MuLV) reverse transcriptase (New England Biolabs) as per manufacturer's instructions. All cDNA was stored at -20°C until required for use in qPCR assays.

3.2.4 Optimization of RT-qPCR Reactions and Confirmation of Controls

Candidate reference gene primers were tested for specificity via the production of a PCR amplicon and subsequent agarose gel electophoresis analysis to confirm fragment sizes. PCR was conducted using GoTaq Green MasterMix (Promega) using the manufacturer's recommended protocol and amplification products were run on a 1% (w/v) agarose-TAE gel (data not shown). An assessment of primer specificity with respect to splicing variants as required by the MIQE guidelines[132] was not required as haloarchaeal

genes do not contain introns and alternative splicing does not occur in any archaeal species. The gradient feature on an iQ5 real-time detection system (BioRad) was used to determine the optimal primer annealing temperature (T_a) for each primer set. Individual reaction conditions (50% (v/v) 2x iQ™ SYBR® Green Supermix (Bio-Rad), 1uM forward primer, 1uM reverse primer) and quantity of template used were held constant for all qPCR assays. The T_a that produced the lowest obtainable quantification cycle (C_q) value was used as the optimized T_a for each specific primer set. No Reverse-Transcription (No-RT) and No-Template (NTC) controls were included in each run to monitor for DNA contamination and to assess the production of primer-dimers[172] via a melt curve conducted on the real-time detection system immediately after the qPCR cycle.

3.2.5 Assessment of RT-qPCR Efficiency

The dynamic linear range (DLR) and amplification efficiency was assessed through the construction of a relative standard curve as described by Taylor *et al*[172]. cDNA from each reverse-transcription reaction above was pooled (250ng from each reaction) and a 1/10 dilution series was created to obtain cDNA standards. qPCR assays (standard recipe, above) were carried out on the pooled cDNA for each serial dilution (100ng, 10ng, 1ng, 0.1ng, and 0.01ng per reaction) using the optimized primer annealing temperatures. The iQ5 software package (Bio-Rad) was used to create standard curves and assess assay efficiency for each primer pair. Curves were required to produce efficiencies of 90-110% and have a curve fit of $R^2 = 0.98$ to be considered acceptable as per the MIQE Guidelines[132].

3.2.6 Analysis of Candidate Reference Gene Stability

The cDNA (10ng; obtained from the linear dynamic range of the standard curves) from each biological replicate grown under each condition was used as template in qPCR assays and ran on the iQ5 real-time detection system described in section 3.2.4 above. Assays were run in technical triplicates (sets of three qPCR reactions constructed from the same sample for the purpose of averaging variation in results produced by the assay) using the optimized primer conditions determined above. Technical triplicates producing a standard deviation of 0.5 or lower were considered acceptable and were included in expression comparisons. A relative-fold change in expression was established by comparing candidate reference gene expression during growth under each test condition to the average expression observed for that gene across all conditions. This was done using the "relative gene expression" function found in the qBase^{plus} software suite (Biogazelle)[173]. The same software was then used to assess reference gene stability using the GeNorm *M*-value as calculated using the GeNorm algarithm[174]. Genes producing an *M*-value of 0.5 or lower were considered as stably expressed.

3.3 Results

Initial assessment of RNA purity via gel electrophoresis identified substantial genomic DNA contamination within RNA extracts. Digestion with DNase I immediately after RNA extraction resulted in RNA extracts of substantially higher purity. Figure 3.1 contains a representative gel used for the assessment of RNA purity ran on a 1% agarose-TAE gel after denaturing formaldehyde agarose gels (ran by standard protocols; data not shown) consistently yielded poor resolution of RNA fragments.

RNA integrity was assessed using BioRad's Experion Mircro-Capillary Electrophoresis system (Figure 3.2). This method utilizes a laser to detect RNA fragments that have been ran out on micro-scale gel within a specialized chip and produces a high-resolution digital image of the fragmentation pattern. This is the preferred method of evaluating RNA integrity stated by the MIQE Guidelines[132] as it is A) highly sensitive; and b) produces a relative quality index (RQI) value based on the densitometric ratio between the 23S and 16S rRNA fragments. The MIQE Guidelines states a minimum RQI of 7 is acceptable; however, all RNA samples used in this study produced an RQI of 9 or greater.

Assessment of candidate reference gene stability began with the optimization of primers (Table 3.3) for the *rpoB* gene as described above. The *rpoB* qPCR assay efficiency (E = 106.6%; R^2 = 0.998) was assessed via cDNA standard curve and identified an optimal cDNA concentration of 10ng per qPCR reaction (see Figure 3.3). The curve was constructed using only 4 of the 5 available serial dilutions due to a consistent lack of fluorescence in the 100ng reaction replicates. This apparent lack of amplification was observed across multiple assays at this cDNA concentration leading to the data point being eliminated from all future standard curves. Further evaluation of these non-fluorescing reactions by agarose gel electrophoresis (data not shown) revealed amplification was in fact occurring as a single fragment of expected size was observed upon visualization with ethidium bromide. Data points were only used in curve construction if the standard deviation between the individual replicates of a given cDNA concentration was 0.5 or lower.

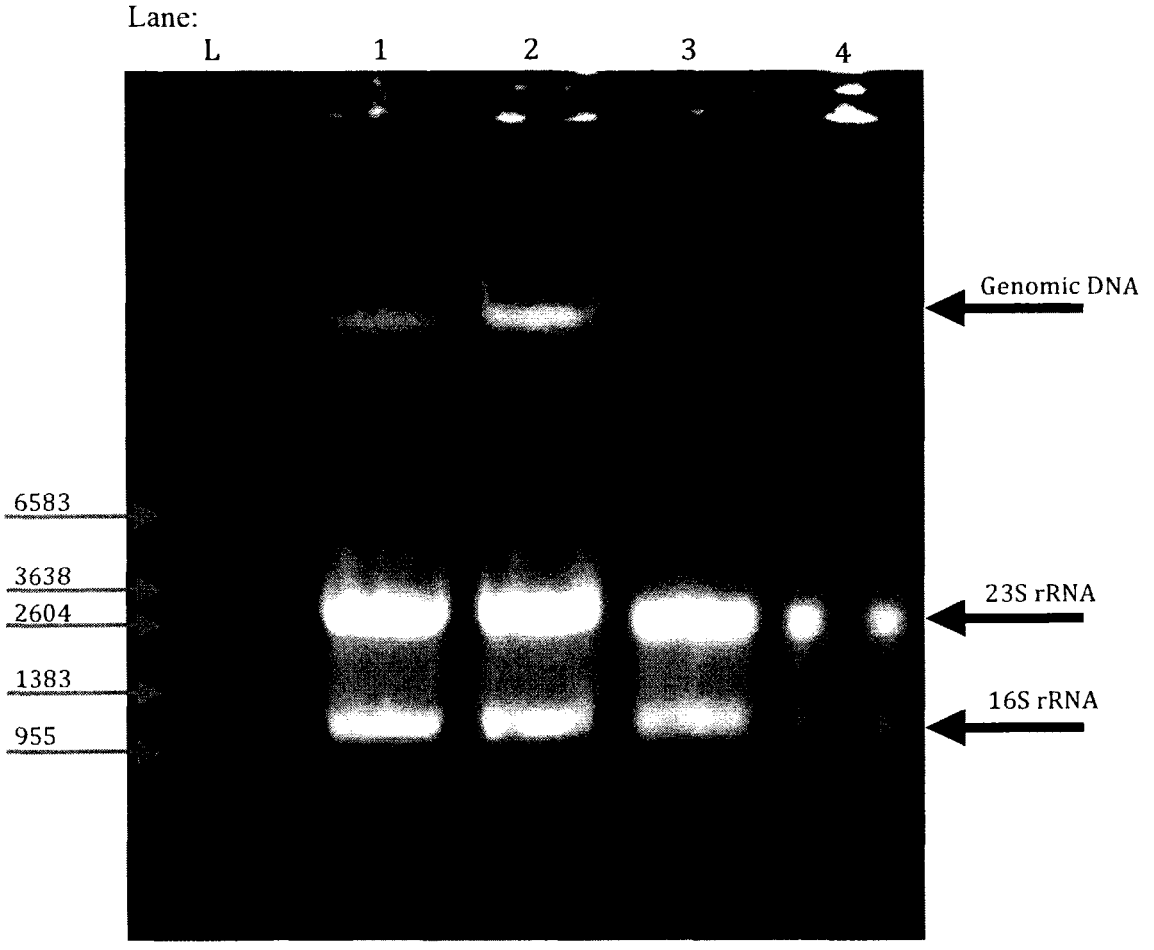

Figure 3.1. Representative RNA purity assessment gel obtained by running RNA extracted from two biological replicates of *Har. marismortui* cells grown in media containing 720mM KCl prior to, and after, digestion with DNase 1. Lanes: L-RNA Ladder; 1-720mM KCl(1), untreated; 2-720mM KCl(2), untreated; 3-720mM KCl(1), DNase digested; 4-720mM KCl(2), DNase digested. Numbers in parentheses behind experimental test conditions indicate biological replicate number. The disruption observed in the fragments in lane 4 was caused by a solid piece of agarose located in the gel immediately in front of the well.

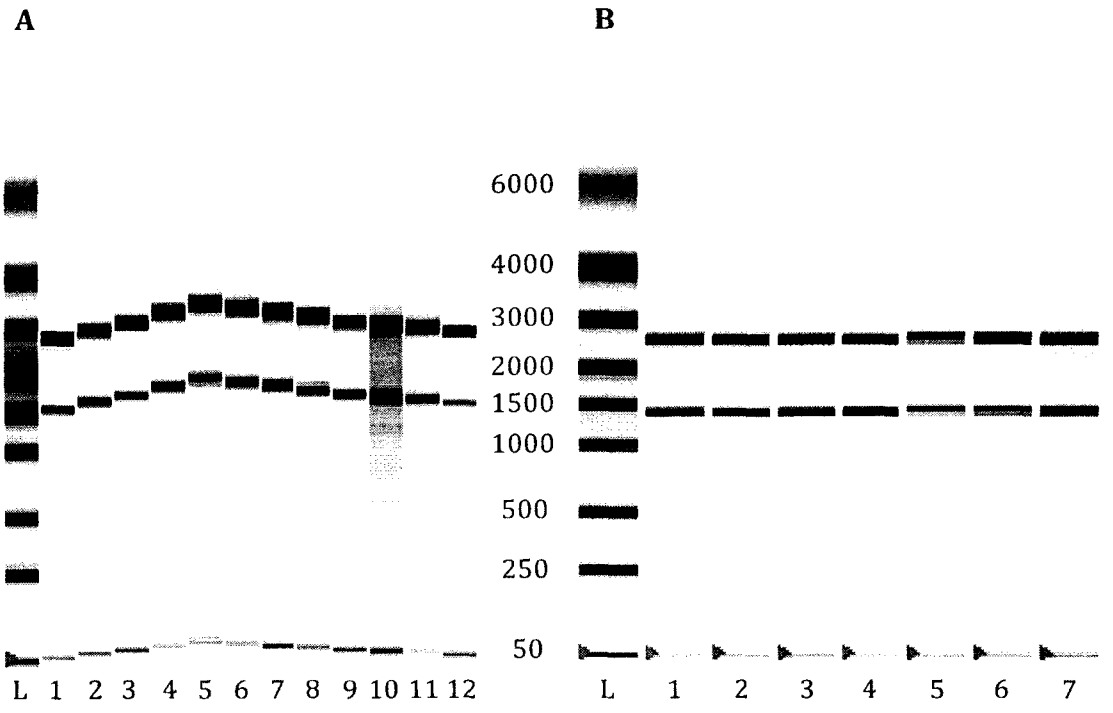

Figure 3.2. Micro-Capillary Electrophoresis gels used for the evaluation of RNA integrity as produced by BioRad's Experion system using the manufacturer's suggested protocol. The 50bp marker is added to each sample as part of the protocol and is used by the software to properly align all lanes to the RNA ladder. Lanes in **A** were not aligned by the software due to the poor resolution of the ladder. All extracted RNA samples used in **B** produced an RQI value of 9 or higher (MIQE recommended RQI = 7). All extracted RNA samples used in **A**, with the exception of the 20mM KCl(1) sample, produced crisp fragments with little apparent degradation and were thus assumed to have an RQI value of 7 or greater. The 20mM KCl(1) sample was re-run in **B** to confirm integrity. The 720mM KCl triplicate was ran twice to confirm reproducibility. **A:** Lanes: L – Ladder; 1 – 45°C(1); 2 - 45°C(2); 3 - 45°C(3); 4 - 37°C(1); 5 - 37°C(2); 6 - 37°C(3); 7 - 55°C(1); 8 - 55°C(2); 9 - 55°C(3); 10 – 20mM KCl(1); 11 – 20mM KCl(2); 12 – 20mM KCl(3). **B:** Lanes: L – Ladder; 1 – 720mM KCl(1); 2 – 720mM KCl(2); 3 – 720mM KCl(3); 4 – 720mM KCl(1); 5 – 720mM KCl(2); 6 – 720mM KCl(3); 7 – 20mM KCl(1). Numbers in parentheses behind experimental conditions indicate biological replicate number.

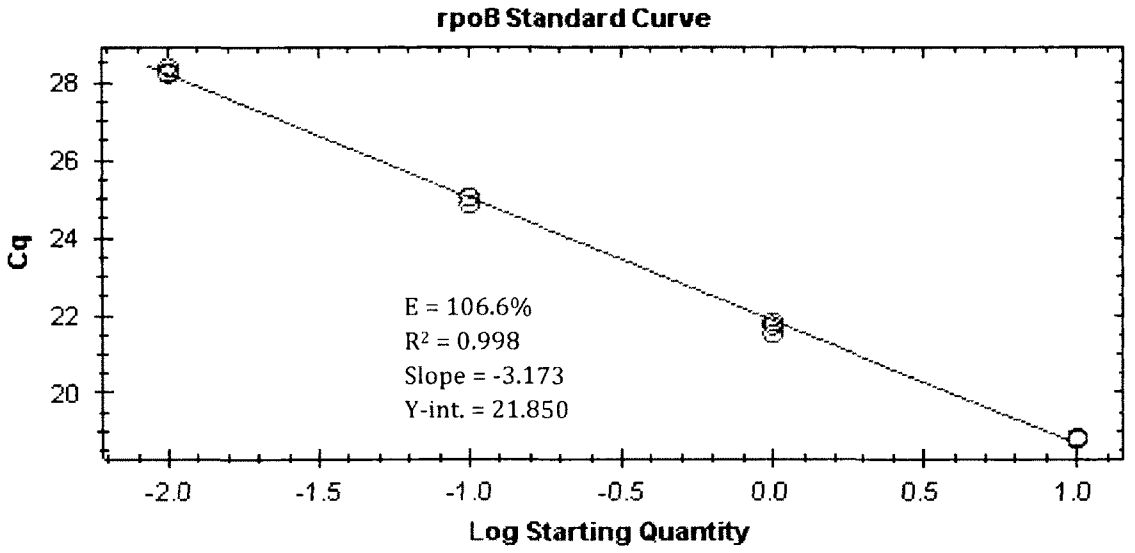

Figure 3.3. Representative RT-qPCR standard curve constructed using the RNA polymerase beta sub-unit (*rpoB*) gene primers. Standards were prepared by pooling cDNA synthesized from RNA extracted from *Har. marismortui* cells under each test condition. Assays were run using a 1/10 serial dilution of pooled cDNA template standards ranging from 10ng to 0.01ng of cDNA per assay reaction covering three logarithmic steps as prescribed by the MIQE Guidelines. Standard deviations above 0.5 were considered unacceptable.

A preliminary assessment of differential gene expression in the *rpoB* gene using two biological replicates under each test condition showed the expression of this gene was highly variable between the two replicates. The assay was repeated two additional times in order to obtain a triplicate of data. Figure 3.4 shows the relative expression change obtained from each of the three individual assays which appear to produce a consistent result. Standard growth conditions were considered normal expression and were assigned a value of 1 thus allowing for a premature relative-fold change to be assigned to alternative expression patterns. Direct comparison of the three assays (Fig 3.4) shows the results for each independent replicate were highly consistent even though substantial variation between biological replicates was observed.

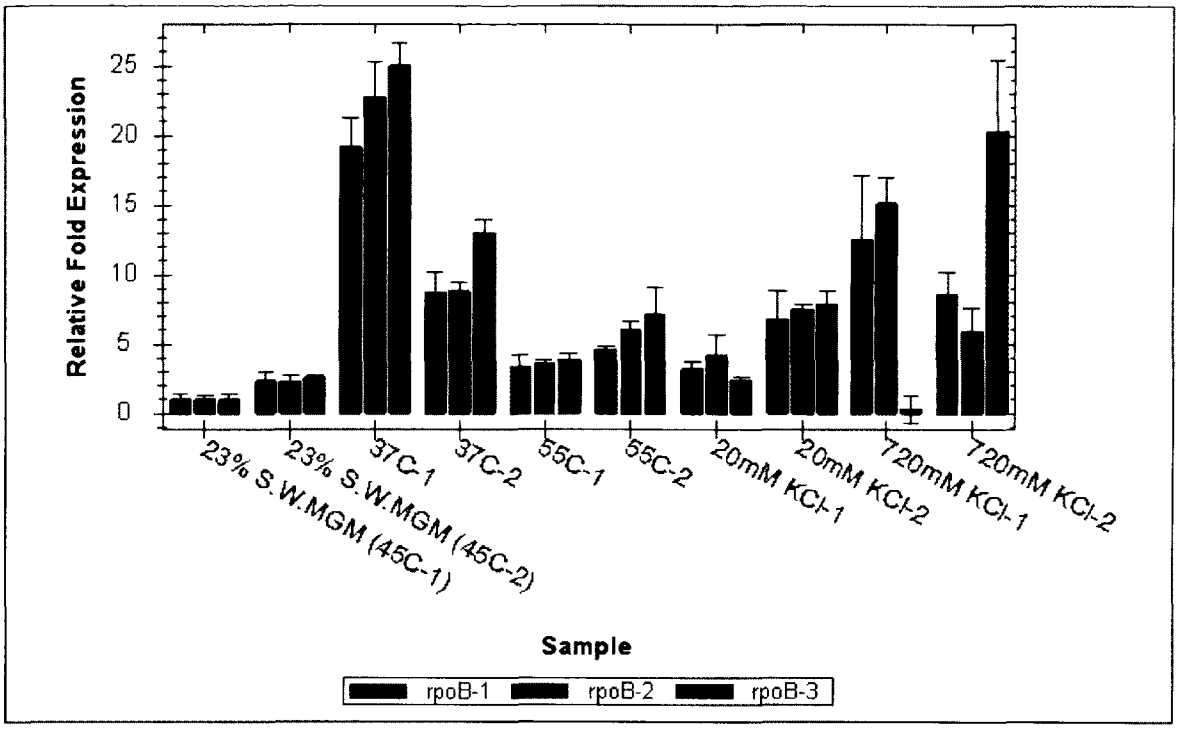

Figure 3.4. Differential expression of the *Har. marismortui* RNA polymerase beta sub-unit gene (*rpoB*) across a range of extracellular potassium concentrations and growth temperatures A direct comparison of three replicate RT-qPCR assays is shown. Replicate assays were conducted on the same samples to assess if the observed variation between assays is due to assay preparation or performance. Expression is shown as fold change relative to the first biological replicate under standard conditions (23% S.W. MGM; 45°C-1). X-axis labels indicate first (-1) and second (-2) biological replicate for each test condition. Each bar is representative of an average result obtained from a technical triplicate within each independent assay. Technical triplicates producing a standard deviation below 0.5 were considered acceptable. **Note:** The drastic change in expression between the 720mM KCl biological replicates (rpoB-3) is due to an error in assay preparation and not a change in differential expression between assays.

A further comparison was conducted in order to assess variation between cDNA synthesis reactions as previous reports have suggested reverse-transcription being primed by random hexamers can substantially over estimate expression levels[162]. As a means of assessing variation between reverse-transcription reactions, three separate cDNA syntheses were conducted from a single extracted RNA sample obtained from a third biological replicate grown under standard conditions (23% S.W. MGM; 45°C). A subsequent RT-qPCR assay was conducted to assess differential expression between the three cDNA samples as a means of determining if the observed variability in the differential expression

between biological replicates is due to natural biological variability or due to the use of random hexamers during priming of the reverse-transcription step. Differential expression between the three cDNA synthesis reactions is shown as a fold change in expression relative to zero in Figure 3.5.

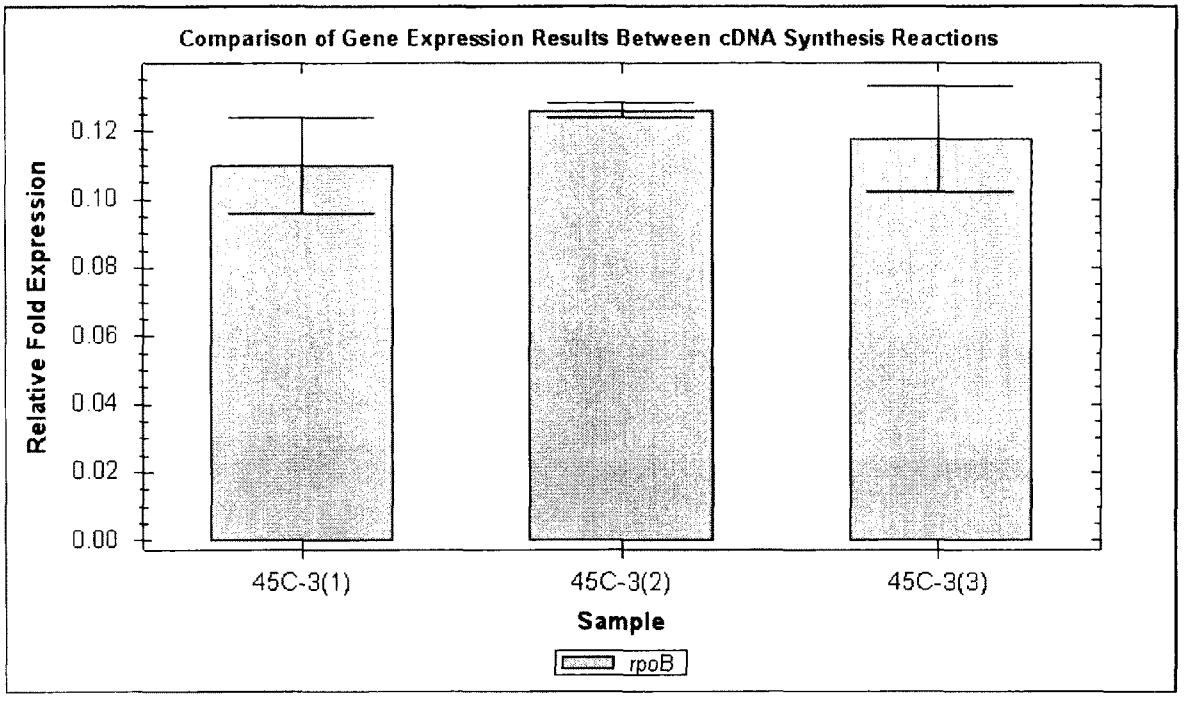

Figure 3.5. Comparison of RNA polymerase beta sub-unit (*rpoB*) expression obtained from three independent cDNA syntheses using a single RNA sample as a template. RNA was extracted from a third biological replicate of *Har. marismortui* cells grown under standard conditions (23% S.W. MGM; 45°C). RNA was extracted after cells were in balanced growth. cDNA synthesis was conducted using random hexameric primers and M-MuLV RT (New England Biolabs) as per manufacturer's recommended protocol. Relative fold expression between cDNA samples is shown relative to 0.

Once a reasonable explanation for the observed variation in gene expression data between biological replicates had been obtained, the stability of five candidate reference genes (16S rRNA, *rpoA*, *rpoB*, *polA2*, and *pykA*) were assessed. RNA was extracted from three newly established biological replicates and cDNA was synthesized as before. Expression data was imported into the qBase^{plus} software suite[173] and used to determine the expression of each candidate reference gene under each condition relative to the average

expression for that gene (Figure 3.6). The software was also used to calculate the GeNorm *M*-value which is the established acceptable statistical measure of expression stability [173] (Figure 3.7).

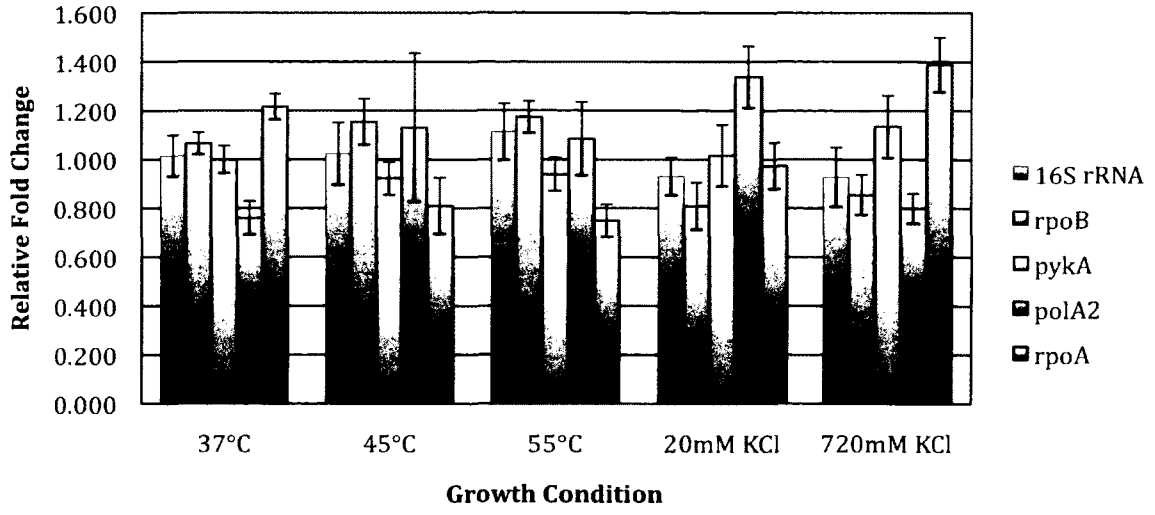

Figure 3.6. Expression of each candidate reference genes under each test condition relative to the average expression of that gene across all test condition. Figure was produced using relative expression values produced by the qBase^{plus} [173] software suite.

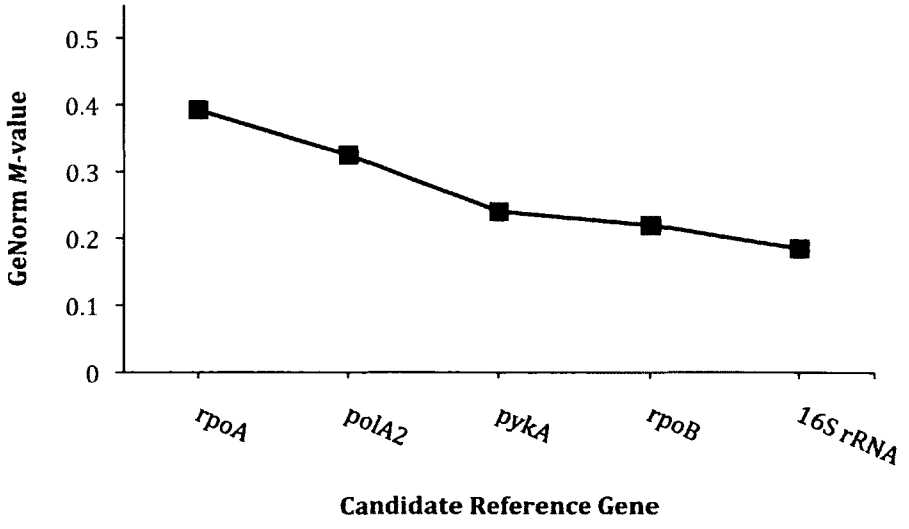

Figure 3.7. Graphical representation of the GeNorm *M*-values obtained using the qBase^{plus} software package. Genes producing values of 0.5 or lower are considered to be stably expressed and may be used as reference genes in qPCR assays as per the MIQE Guidelines.

3.4 Discussion

Although the standard curve produced for the RNA polymerase candidate reference gene meets the requirements of the MIQE guidelines[132] in terms of the number of logarithmic steps used to construct the curve, it would have been ideal to construct the curve using all 5 planned cDNA standards. Curiously, the expected increase in fluorescence of the qPCR reactions containing 100ng of cDNA did not rise above the base-line level of fluorescence. Since auto-quenching has not been observed in any assay utilizing SYBR Green to the best of our knowledge, the most logical answer to this phenomenon would simply be that insufficient reagents (dNTPs, primers, etc.) were available for cDNA replication to occur. Intuitively, however, this seems unlikely considering the quantity of primer and dNTPs used in each 25μL assay are typical of a 50μL reaction which would be sufficient to amplify several hundred nanograms of template. Moreover, non-specific primer annealing leading to random amplicon production or highly inefficient amplification is also unlikely as all reactions carried out for each standard curve were prepared from a single master mix of reagents. Had primer specificity caused this issue all reactions used in the construction of the standard curve would have been non-functional. As per the MIQE Guidelines[132], melt curve analysis was conducted alongside each assay and did not produce results that suggest non-specific PCR products had been produced. This consistently repeatable lack of fluorescence in the 100ng cDNA standards lead to these data points being eliminated from standard curve construction all together. Agarose gel electrophoresis indicated amplification was in fact occurring normally as single fragments of anticipated size were observed suggesting fluorophore quenching is occurring as a result of a specific reaction condition or potentially due to auto-quenching of the SYBR Green fluorophore.

Preliminary quantification of the *rpoB* gene via a triplicate of RT-qPCR assays revealed consistency of expression for each test condition within a single biological replicate. When gene expression under each unique test condition was compared to a second independent biological replicate under the same condition, substantial variance was observed (see Figure 3.4). Considering that biological replicate cultures were grown in parallel to one another using media from the same stock and under identical temperature conditions, the level of gene expression between two biological replicates would be expected to be identical. Seeing as this was not the case, it is possible that the observed variability between biological replicates is due to the method of priming used for cDNA syntheses. Since Lopez-Lopez *et al*[162] have shown that the use of random hexameric primers can result in an overestimation of mRNA copy number by up to 20-fold, the biological variation being observed could potentially be due to how the mRNAs from each biological replicate was primed for cDNA synthesis. The observed relative fold change of mRNA copy number could show substantial variation between identically treated replicates simply by changing the position or sequence, of an annealing hexameric primer. Variation in the observed level of expression between biological replicates may also occur if a single template is being primed a variable number of times within a reverse-transcription reaction thus making the reaction less linear then it is intended to be. Because either of these scenarios could be occurring, it was necessary to evaluate candidate reference gene stability based on an average value obtained with the inclusion of additional biological replicates.

After considering the scenarios above, the potential variation between independent reverse-transcription reactions that utilize a common RNA template was assessed. Three independent cDNA samples were obtained from a third biological replicate of *Har. marismortui* cells grown under standard conditions and expression of the *rpoB* gene was

assessed for each. As observed in Figure 3.5, minimal variation in gene expression is observed between the cDNA samples. Since the relative fold change in expression between the samples fall within an acceptable standard deviation of one another, and because the maximum fold change relative to zero is 0.12, we can assume the variation observed in figures 3.4 can be attributed to natural biological variability.

Assessment of reference gene stability using the qBase^{plus} software suite[173] has revealed that all five of the candidate reference genes selected initially can be used as reference genes under the conditions used above. Through examination of the relative fold change data observed (Figure 3.6) it is apparent that the *rpoA* and *polA2* genes are the least stably expressed as the expression of these genes is shown to decrease by approximately 25% under some conditions and increase by approximately 40% under others. This observation is supported by the GeNorm *M*-values produced by the qBase^{plus} software[173]. As per the MIQE Guidelines[132], genes producing GeNorm *M*-values below 0.5 are considered as stably expressed and are appropriate for use as a reference. The GeNorm *M*-values of *rpoA* and *polA2* are both below 0.5 suggesting they are appropriate reference genes; however, these are considerably higher then the remaining three candidates which all have GeNorm *M*-values of 0.25 or lower. Placing the five candidates in order of increasing GeNorm *M*-value ranks these genes by increasing suitability as qPCR reference genes. This makes the total expression of all 16S rRNA genes (determined using the primers in Table 3.2 which amplify a homologous region from all 16S rRNA operons) the most ideal reference gene for qPCR studies followed by the *rpoB*, *pykA*, *polA2*, then *rpoA* genes. The *pykA* gene, encoding pyruvate kinase has, to best of our knowledge, never before been validated as a reference gene for qPCR studies making it an entirely novel control.

In the event any of the five novel reference genes that have now been properly established are determined to be unacceptable for use as a reference gene under a future growth condition of interest, a second subset of candidates (Table 3.3) has been selected using whole genome microarray data from a closely related halophilic archaeon, *Halobacterium salinarum* NRC-1. These data are openly available through the online Gaggle database and software package[175-177] and contains differential gene expression data obtained by microarray analysis for all 2400 genes in the NRC-1 genome across 361 independent growth conditions. The $\log_{10}$ ratios for each gene provided by the database have been used to determine the coefficient of variance across all available experimental conditions (data not shown). The genes producing the lowest overall coefficients have been selected as candidate reference genes and RT-qPCR primers have been designed using the Beacon Designer 7 software suite as described above. Interestingly, when the NRC-1 homologues of the five genes evaluated in this study are re-evaluated using this microarray data they all fall just beyond the coefficient of variance cut-off that was used to create the second subset of candidates found in Table 3.3. This suggests that the new subset of candidates are very likely to be verified as highly appropriate reference genes and may set a precedent for determining candidate reference genes based on expression levels observed via microarray analysis.

Table 3.3. Candidate reference genes identified as being minimally variable by *Halobacterium salinarum* NRC-1 micro array data obtained from the Gaggle database[175-177].

Gene	NRC-1 Loci	*Haloarcula* Loci*	*Haloarcula* Gene Product
sun	VNG0499G	rrnAC0628	Cytosine-C5-methylase
-	VNG0508H	rrnAC0928	Hypothetical Protein
-	VNG1514H	rrnAC0981	Hypothetical Protein
SucC	VNG1541G	rrnAC0472	Succinyl-CoA synthetase beta chain
SucD	VNG1542G	rrnAC0474	Succinyl-CoA synthetase alpha chain
zim	VNG1543G	rrnAC1876	CTAG modification methylase
CbiT	VNG1550G	rrnAC2998	Precorrin-8W decarboxylase
CbiG	VNG1555G	rrnAC3003	Cobalamin biosynthesis protein
CbiH	VNG1557G	rrnAC3008	Precorrin-3B C17-methyltransferase
-	VNG1558H	rrnAC3009	Hypothetical Protein
-	VNG1559H	rrnAC3010	Hypothetical Protein
CbiK	VNG1561C	rrnAC3011	Cobalt chelase thioredoxin
-	VNG1562H	rrnAC3012	Hypothetical Protein
Tup1	VNG1564H	rrnAC3013	*Tup1*-like transcriptional repressor
CobN	VNG1566G	rrnAC3019	Cobalamin biosynthesis protein
CbiJ	VNG1568G	rrnAC3021	Cobalt-precorrin-6YC5-methyltransferase
etfB1	VNG2150G	rrnAC3148	Electron transfer flavoprotein beta sub-unit
dpa	VNG2462G	rrnAC3118	Signal recognition particle receptor

*These loci were identified as *Haloarcula marismortui* homologues to the provided NRC-1 loci using the BLAST feature on the UCSC Archaeal Genome Browser[85, 86, 170, 171].

3.5 Conclusion

In summary, we have shown that investigations utilizing RT-qPCR as a tool for studying gene expression in archaeal organisms can conform to the 2009 MIQE Guidelines[132] with relatively minimal additional effort and have established an approach to doing so that can be mimicked by other research groups if required. The use of the increasingly common RNA spin column as a quick and effective method of RNA extraction from halophilic cells has proven to yield not only large quantities of RNA, but RNA of

extremely high integrity as shown by the RQI values obtained from the Experion Micro-Capillary Electrophoresis system. This was a point of initial concern due to the presence of large salt concentrations that could have potentially affected the spin column membrane and thus the efficiency of the RNA extraction. A preliminary evaluation of differential expression of the *rpoB* gene, which appears to be highly variable when grown across a range of extracellular potassium ion concentrations and growth temperatures, has shown that the methodology of using random hexameric primers for the single strand synthesis of cDNA is highly consistent leaving the variation in cellular RNA levels for this gene to be attributed to biological variability between cellular replicates. Although these methodologies may need to be adjusted to accommodate the extraction of RNA from other archaeal species due to the unique environments they live in (For example, how well will an RNA spin column work when isolating cells from media at a pH of 2?), the techniques described above create a definitive starting point for the evaluation of reference genes in many other organisms.

We have now identified and validated five novel reference genes in the halophilic archaeon, *Haloarcula marismortui*, which we believe is a valuable starting point for the validation of proper reference genes in many other archaeal species. Although each of the genes evaluated above have met a minimum set of standards for use as RT-qPCR references, their expression should still be compared to the expression of genes that would be expected to be highly variable under the growth conditions used to further confirm stability. This will be completed in the very near future in order to provide additional support for these novel reference genes, whether it is truly required or not. Additionally, a precedent has now been established for the use of largely available microarray expression data to select potential RT-qPCR reference gene candidates. This approach has now been

used to identify a second suitable subset of candidate haloarchaeal reference genes that can be utilized by other research groups to evaluate and validate should the novel reference genes above prove to be unsuitable for a specific growth condition. It is our hope that this investigation will introduce the MIQE guidelines[132] to the archaeal research community and demonstrate that these guidelines can, and should be, applied to RT-qPCR studies that are currently lacking standardization of experimental design and methodology.

Chapter Four

Conclusion

4.1 Assessment of Potassium Stress Responses in *Haloarcula marismortui*

Substantial progress toward better understanding how the halophilic archaeon, *Haloarcula marismortui*, responds to potassium stresses in its native environment has now been made. This has been completed by establishing cellular growth characteristics under standard laboratory growth conditions then evaluating changes in these characteristics under extremely elevated and limiting concentrations of extracellular potassium. This investigation has shown that *Har. marismortui* exhibits an optimal generation time of 4.19 ± 0.14 hours at an extracellular K^+ concentration of 100mM. The cellular generation time increases substantially as K^+ reaches limiting concentrations (8mM) but only slightly as extracellular K^+ increases to concentrations nearly an order of magnitude higher (720mM) than optimal. We have suggested that the increased generation time observed in *Har. marismortui* as K^+ becomes limiting is very likely a result of changing energy requirements within the cell. Because *Har. marismortui* sequesters K^+ to intracellular concentrations of nearly 1.4M when the extracellular concentration is 8mM (Chapter 2; Figure 2.4), an extremely steep ion gradient is formed that is likely to result in efflux of K^+ through low-affinity potassium channels in the cellular membrane such as the low-affinity Trk potassium transport system. This efflux should result in the immobilization of ATP as a regulator of the Trk system in order to maintain the critical intracellular K^+ concentration at adequate levels which, in turn, would result in a reduced energy reserve for normal cellular system function. A reduced energy capacity for normal cell function will result in slowed generation times.

Alternatively, as extracellular K^+ is elevated in excess of 700% above normal, the cellular generation time increases only slightly as intracellular K^+ concentrations approach 2.5M (Chapter 2; Figure 2.4) which may suggest a potential inability to sufficiently regulate K^+ uptake. If an uncontrollable uptake of K^+ is in fact occurring, an accumulation of intracellular salts to concentrations that become sub-optimal for cell system function is likely to cause a deterioration in generation time. As other members of our lab group are demonstrating with currently unpublished data, the proper function of many halophilic enzymes are drastically affected by elevated K^+ concentrations *in vitro* thus making this a plausible explanation for the observed increases in generation time.

In addition to the valuable information obtained from the aforementioned growth studies, we have now established a methodology for using modern analytical chemistry instrumentation in the determination of intracellular ion concentrations. We have shown this methodology to be highly ineffective when examining ions of lower molecular weight such as sodium and lithium. This is due to the nature of Induction-Coupled Plasma instrumentation and the space-charge effects that occur when high concentrations of low-molecular weight ions are introduced into argon plasma. Though this is the case for some ions that we had hoped to evaluate, it appears to produce reliable, or at minimum plausible, data when examining the concentrations of higher molecular weight ions. Using this methodology we have shown that *Har. marismortui* sequesters K^+ ions to concentrations several fold higher than that of the media it grows in and this uptake increases or decreases with the extracellular concentration of the ion. This is presumably done in order to balance the osmotic pressure experienced in its native environment as has been previously suggested[112]. We have shown that the organism is an excellent scavenger of K^+ ions under limiting conditions and will readily uptake Rb^+, as has been previously reported.

Interestingly, we have also demonstrated that *Har. marismortui* will uptake Cs^+ in place of K^+ if these ions are available; however, the use of alternative monovalent ions as a replacement for K^+ results in extremely slowed, and apparently variable generation times. Additionally, these ions are not sequestered to concentrations observed for potassium under optimal conditions which may further suggest that K^+ is required to maintain cellular energetics in a manner that has not yet been described.

4.2 Future Direction for the Evaluation of Potassium Stress Responses in Halophilic Archaea

Now that we have identified and validated several RT-qPCR reference genes (Chapter 3 and Section 4.3) a secondary 'test' study will be initiated to examine the differential expression of the gene products within the Trk system described in *Haloferax volcanii*[137]. Several genes of particular interest are those that comprise the Trk low-affinity potassium transport system. Several of these genes are known to encode membrane bound potassium ion transport proteins[145, 178] which would be expected to demonstrate highly differential expression patterns under the conditions used to verify our five novel reference genes. Previous studies have shown that the gene products of the Trk system genes will actively transport rubidium as a substitute to potassium[145] in some species making it of particular interest with regards to the findings described in Chapter 2. Once gene specific RT-qPCR primers have been optimized and assay efficiency has been assessed in accordance with the MIQE guidelines, the differential expression of the many genes that comprise the Trk system in *Har. marismortui* will be examined. Differential expression will be determined across a range of potassium and rubidium ion concentrations in the growth media while using the reference genes identified in this study as endogenous controls.

Reference gene stability will also be re-evaluated, as per the MIQE Guidelines[132], using the methodologies outlined to ensure stability is not affected by growth in rubidium. This secondary study will further demonstrate the usefulness of our reference genes and further our knowledge of *Haloarcula marismortui* responses to potassium stress.

In addition to the latter planned investigation, it would be of interest to monitor changes in potassium transporter activity in *Har. marismortui* under *in vivo* conditions across our chosen range of extracellular K^+ concentrations. This can be done using an assay such as the FluxOR™ Potassium Ion Channel Assay (Invitrogen) which utilizes the well documented affinity that potassium channels exhibit for Thallium ions. The assay works by introducing thallium ions into the cell via flow along their concentration gradient or via natural active transport. The ions then interact with, and stimulate a membrane permeable reporter dye resulting in cytosolic fluorescence. The intensity of observed fluorescence is then directly proportional to the total activity of all potassium ion channels present within the cell. This would very likely provide outstanding supporting evidence for any differential expression of ion transporters observed during the RT-qPCR study outlined above. Although this assay is quite costly and the manufacturer's website states it is intended for use with mammalian cells, it would be well worth attempting to adapt this assay for use with haloarchaeal systems.

4.3 Evaluation of Novel RT-qPCR Reference Genes in *Haloarcula marismortui*

We have now shown that the archaeal research community should be able to conform to the MIQE Guidelines[132] while adding only minimal additional effort to established RT-qPCR protocols. The purpose of the 2009 publication by Bustin *et al*[132] was not to inflict excessive amount of additional work on investigators conducting quantitative

nucleic acid analyses, but to ensure that all investigators are following a set of standardized rules that maximum the reliability and reproducibility of data. As this technique can be highly variable due to the potential for many arbitrary threshold and cut off values while analyzing data, and due to the numerous techniques and products available for RNA isolation, purification, and quantification, cDNA synthesis, and the qPCR assay itself, it is important that sufficient information regarding methodology and experimental design be provided in qPCR publications.

After realizing that most investigators using RT-qPCR as a technique for studying archaeal organisms do not appear to be conforming to the MIQE Guidelines due to a severe lack of available reference genes that have been properly validated, we are now prepared to provide the archaeal research community with an established protocol that can be adapted as needed to validate their own candidate reference genes in various archaeal species. Five novel reference genes have now been validated by this methodology in *Har. marismortui*, one of which (*pykA*; pyruvate kinase) has, to the best of our knowledge, never before been used as a reference gene in any species. In order of decreasing expression stability, these genes are *16S rRNA*, *rpoB*, pykA, *polA2*, and *rpoA* which code for the 16S ribosomal RNA, the RNA-polymerase beta subunit, pyruvate kinase, the DNA-polymerase II large subunit, and the RNA-polymerase alpha subunit, respectively. We have also shown that existing gene expression data obtained from microarray analysis can be successfully utilized for the initial identification of reference gene candidates upon initiation of a reference gene search.

In addition to the validation of our novel reference genes, preliminary studies have been used to show that substantial biological variation in gene expression is observed between independent biological replicates of *Haloarcula marismortui*; a phenomenon that is likely to be observed in other archaeal species. After initially attributing this variation to

the method of priming used for the cDNA synthesis reaction, the use of random hexameric primers has been shown to produce acceptably consistent results upon comparison of expression data obtained from multiple cDNA samples synthesized independently from the same RNA extract. This is of substantial importance as most archaeal species do not possess poly-adenylated tails on the RNAs eliminating the possibility of using oligo-dT primers as a cDNA priming method (see Section 3.1.1). Although gene specific primers can be used for cDNA synthesis, this method can become extremely laborious and costly when high-throughput is required while evaluating reference genes. As stated in Chapter 3, it is our hope that this investigation will introduce the MIQE guidelines[132] to the archaeal research community and demonstrate that these guidelines can, and should be, applied to RT-qPCR studies that are currently lacking standardization of experimental design and methodology.

4.4 Further Evaluation of Identification of Novel RT-qPCR Reference Genes in Halophilic Archaea

As mentioned in Section 4.2, the five novel reference genes identified in this study will now be used to assess the differential expression of several genes within the low-affinity Trk potassium transport system. This will be part of an ongoing goal to unequivocally validate a number of useable reference genes in *Haloarcula marismortui*. The differential expression data for the Trk protein genes will be evaluated as per the methodology used above. The stability of these genes will then be compared to that of the five novel reference genes to further show that the stability of our validated references is not an artifact of an insufficient dataset.

Additionally, the evaluation of several of the subset 2 candidate reference genes (Table 3.3) in *Har. marismortui* is nearly underway. The validation of several more genes will allow us to produce a wide range of acceptable haloarchaeal RT-qPCR reference genes that can be readily used or re-evaluated under new growth conditions by other investigators. This will also provide several "go-to" alternatives in the event one or more validated genes are shown to exhibit reduced stability under one or more specific growth conditions.

Finally, all novel reference genes that are validated in *Har. marismortui*, current and future, will also be evaluated in the closely related halophilic species, *Halobacterium salinarum* species NRC-1. This would seem appropriate as all of candidate genes that are about to be evaluated in *Har. marismortui* were identified using the NRC-1 microarray data obtained from the Gaggle database and software package[175-177]. This will not only allow for "proof of concept", but may identify reference genes that are universal among the haloarchaea.

4.5 References

1. Stanier, R.Y. & van Niel, C.B. The concept of a bacterium. *Archiv fur Mikrobiologie* **42**, 17-35 (1962).

2. Sapp, J. The prokaryote-eukaryote dichotomy: meanings in mythology. *Microbiology and Molecular Biology Reviews* **69**, 292-305 (2005).

3. Chatton, E. Titres et Travaux Scientifiques (1906-1937) de Edouard Chatton. (Sette, Sottano, Italy; 1938).

4. Haeckel, E. Generelle morphologie der organismen. (Reimer, G., 1866).

5. Haeckel, E. The Wonders of Life: A Popular Study of Biological Philosophy. (Harper and Brothers, New York, NY; 1904).

6. Copeland, E.B. What is a plant? *Science* **65**, 388-390 (1927).

7. Copeland, H.F. The kingdoms of organisms. *The Quarterly Review of Biology* **13**, 383-420 (1938).

8. Copeland, H.F. The Classification of Lower Organisms. (Pacific Books, Palo Alto, CA; 1956).

9. Whittaker, R.H. On the broad classification of organisms. *The Quarterly Review of Biology* **34**, 210-226 (1959).

10. Whittaker, R.H. New concepts of kingdoms of organisms. *Science* **163**, 150-160 (1969).

11. Sogin, S.J., Sogin, M.L. & Woese, C.R. Phylogenetic measurement in procaryotes by primary structural characterization. *Journal of Molecular Evolution* **1**, 173-184 (1972).

12. Woese, C.R., Sogin, M.L. & Sutton, L.A. Procaryote phylogeny. *Journal of Molecular Evolution* **3**, 293-299 (1974).

13. Woese, C.R. in Archaea. Molecular and Cellular Biology. (ed. R. Cavicchioli) 1-13 (American Society for Microbiology Press, Washington, DC; 2007).

14. Fox, G.E. et al. The phylogeny of prokaryotes. *Science* **209**, 457-463 (1980).

15. Sanger, F., Brownlee, G.G. & Berrell, B.G. A two-dimensional fractionation procedure for radioactive nucleotides. *Journal of Molecular Biology* **13**, 373-398 (1965).

16. Woese, C.R. Q & A. *Current Biology* **15**, R111-R112 (2005).

17. Doolittle, W.F., Woese, C.R., Sogin, M.L., Bonen, L. & Stahl, D. Sequence studies on 16S ribosomal RNA from a blue-green alga. *Journal of Molecular Evolution* **4**, 307-315 (1975).

18. Woese, C.R. & Fox, G.E. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences* **74**, 5088-5090 (1977).

19. Zablen, L.B., Kissil, M.S., Woese, C.R. & Buetow, D.E. Phylogenetic origin of the chloroplast and prokaryotic nature of its ribosomal RNA. *Proceedings of the National Academy of Sciences* **72**, 2418-2422 (1975).

20. Balch, W.E., Magrum, L.J., Fox, G.E., Wolfe, R.S. & Woese, C.R. An ancient divergence among bacteria. *Journal of Molecular Evolution* **9**, 305-311 (1977).

21. Fox, G.E., Magrum, L.J., Balch, W.E., Wolfe, R.S. & Woese, C.R. Classification of methanogenic bacteria by 16S ribosomal RNA characterization. *Proceedings of the National Academy of Sciences* **74**, 4537-4541 (1977).

22. Wolfe, R.S. in The Prokaryotes, Edn. 3. (eds. M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer & E. Stackebrandt) 3-9 (Springer-Verlag, New York, NY; 2006).

23. Woese, C.R., Kandler, O. & Wheelis, M.L. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proceedings of the National Academy of Sciences* **87**, 4576-4579 (1990).

24. Woese, C.R., Magrum, L.J. & Fox, G.E. Archaebacteria. *Journal of Molecular Evolution* **11**, 245-252 (1978).

25. Kates, M. in Ether Lipids: Chemistry and Biology. (eds. F. Snyder & W.J. Bauman) 251-398 (Academic Press, New York, NY; 1972).

26. Langworthy, T.A., Smith, P.F. & Mayberry, W.R. Lipids of *Thermoplasma acidophilum*. *Journal of Bacteriology* **112**, 1193-1200 (1972).

27. Langworthy, T.A., Mayberry, W.R. & Smith, P.F. Long-chain glycerol diether and polyol dialkyl glycerol triether lipids of *Sulfolobus acidocaldarius*. *Journal of Bacteriology* **119**, 106-116 (1974).

28. de Rosa, M., Gambacorta, A. & Bu'Lock, J.D. Extremely thermophilic acidophilic bacteria convergent with *Sulfolobus acidocaldarius*. *Journal of General Microbiology* **86**, 156-164 (1975).

29. de Rosa, M., Gambacorta, A. & Bu'Lock, J.D. The Caldariella group of extreme thermoacidophile bacteria: direct comparison of lipids in Sulfolobus, Thermoplasma, and the MT strains. *Phytochemistry* **15**, 143-145 (1976).

30. Boucher, Y. in Archaea. Molecular and Cellular Biology. . (ed. R. Cavicchioli) 341-353 (American Society for Microbiology Press, Washington, DC; 2007).

31. de Rosa, M., Gambacorta, A., Minale, L. & Bu'Lock, J.D. Cyclic diether lipids from very thermophilic acidophilic bacteria. *Chemical Communications*, 543-544 (1974).

32. Jahnke, L.L. et al. Signature Lipids and Stable Carbon Isotope Analyses of Octopus Spring Hyperthermophilic Communities Compared with Those of Aquificales Representatives. *Applied and Environmental Microbiology* **67**, 5179-5189 (2001).

33. Gattinger, A., Schloter, M. & Munch, J.C. Phospholipid etherlipid and phospholipid fatty acid fingerprints in selected euryarchaeotal monocultures for taxonomic profiling. *FEMS Microbiology Letters* **213**, 133-139 (2002).

34. Kaneshiro, S.M. & Clark, D.S. Pressure effects on the comnposition and thermal behaviour of lipids from the deep-sea thermophile *Methanococcus jannaschii*. *Journal of Bacteriology* **117**, 3668-3672 (1995).

35. Sprott, G.D., Meloche, M. & Richards, J.C. Proportions of diether, macrocyclic diether, and tetraether lipids in *Methanococcus jannaschii* grown at different temperatures. *Journal of Bacteriology* **173**, 3907-3910 (1991).

36. Gliozzi, A., Paoli, G., de Rosa, M. & Gambacorta, A. Effect of isoprenoid cyclization on the transition temperature of lipids in thermophilic archaebacteria. *Biochimica et Biophysica Acta* **735**, 234-242 (1983).

37. Nichols, D.S. et al. Cold Adaptation in the Antarctic Archaeon Methanococcoides burtonii Involves Membrane Lipid Unsaturation. *Journal of Bacteriology* **186**, 8508-8515 (2004).

38. Brown, A.D. & Shorey, C.D. Cell envelopes of two extremely halophilic bacteria. *Journal of Cell Biology* **18**, 681-689 (1963).

39. Kushner, D.J. & Onishi, H. Absence of normal cell wall constituents from the outer layers of *Halobacterium cutirubrum*. *Canadian Journal of Biochemistry* **46**, 997-998 (1968).

40. Brown, A.D. & Cho, K.Y. The walls of extremely halophilic cocci: gram-positive bacteria lacking muramic acid. *Journal of General Microbiology* **62**, 267-270 (1970).

41. Darland, G., Brock, T.D., Samsonoff, W. & Conti, S.F. A thermophilic, acidophilic mycoplasma isolated from a coal refuse pile. *Science* **170**, 1416-1418 (1970).

42. Brock, T.D., Brock, K.M., Belly, R.T. & Weiss, R.L. *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Archiv fur Mikrobiologie* **84** (1972).

43. Reistad, R. Cell wall of an extremely halophilic coccus. *Archiv fur Mikrobiologie* **82**, 24-30 (1972).

44. Weiss, R.L. Subunit cell wall of *Sulfolobus acidocaldarius*. *Journal of Bacteriology* **118**, 275-284 (1974).

45. Kandler, O. & Hippe, H. Lack of peptidoglycan in the cell walls of *Methanosarcina barkeri*. *Archives of Microbiology* **113**, 57-60 (1977).

46. Jones, J.B., Bowers, B. & Stadtman, T.C. *Methanococcus vannielii*: ultrastructure and sensitivity to detergents and antibiotics. *Journal of Bacteriology* **130**, 1357-1363 (1977).

47. Konig, H., Rachel, R. & Claus, H. in Archaea. Molecular and Cellular Biology. (ed. R. Cavicchioli) 315-340 (American Society for Microbiology Press, Washington, DC; 2007).

48. Steber, J. & Schleifer, K.H. *Halococcus morrhuae*: a sulfated heteropolysaccharide as the structural component of the bacterial cell wall. *Archives of Microbiology* **105**, 173-177 (1975).

49. Rachel, R., Wyschkony, I., Riehl, S. & Huber, H. The ultrastructure of *Ignicoccus*: evidence for a novel outer membrane and for intracellular vesicle budding in an archaeon. *Archaea* **1**, 9-18 (2002).

50. Zillig, W., Stetter, K.O. & Janekovic, D. DNA-dependent RNA polymerase from the archaebacterium *Sulfolobus acidocaldarius*. *European Journal of Biochemistry* **96**, 597-604 (1979).

51. Langer, D., Hain, J., Thuriaux, P. & Zillig, W. Transcription in Archaea: smiliarity to that in Eucarya. *Proceedings of the National Academy of Sciences* **92**, 5768-5772 (1995).

52. Bartlett, M.S. Determinants of transcription initiation by archaeal RNA polymerase. *Current Opinion in Microbiology* **8**, 677-684 (2005).

53. Hirata, A. & Murakami, K.S. Archaeal RNA polymerase. *Current Opinion in Structural Biology* **19**, 724-731 (2009).

54. Bell, S.D., Jaxel, C., Nadal, M., Kosa, P.F. & Jackson, S.P. Temperature, template topology, and factor requirements of archaeal transcription. *Proceedings of the National Academy of Sciences* **95**, 15218-15222 (1998).

55. Woychik, N.A. & Hampsey, M. The RNA polymerase II machinery: structure illuminates function. *Cell* **108**, 453-463 (2002).

56. Hirata, A., Klein, B.J. & Murakami, K.S. The X-ray crystal structure of RNA polymerase from Archaea. *Nature* **451**, 851-854 (2008).

57. Korkhin, Y. et al. Evolution of complex RNA polymerases: the complete archaeal RNA polymerase structure. *PLoS Biology* **7**, 0001-0010 (2009).

58. Abelson, J., Trotta, C.R. & Li, H. tRNA splicing. *Journal of Biological Chemistry* **273**, 12685-12688 (1998).

59. Marck, C. Identification of BHB splicing motifs in intron-containing tRNAs from 18 archaea: evolutionary implications. *RNA* **9**, 1516-1531 (2003).

60. Watanabe, Y.-i. et al. Introns in protein-coding genes in archaea. . *FEBS Letters* **510**, 27-30 (2002).

61. Thompson, L.D. & CDaniels, C.J. Recognition of exon-intron boundaries by the *Halobacterium volcanii* tRNA intron endonuclease. *Journal of Biological Chemistry* **265**, 18104-18111 (1990).

62. Woese, C.R., Gupta, R., Hahn, C., M., Zillig, W. & Tu, J. The phylogenetic relationships of three sulfur dependent archaebacteria. *Systematic and Applied Microbiology* **5**, 97-105 (1984).

63. Chan, P.P., Cozen, A.E. & Lowe, T.M. Discovery of permuted and recently split transfer RNAs in Archaea. *Genome Biology* **12**, R38 (2011).

64. Rozenski, J., Crain, P.F. & McCloskey, J.A. The RNA modification database: 1999 update. *Nucleic Acids Research* **27**, 196-197 (1999).

65. Noon, K.R., Bruenger, E. & McCloskey, J.A. Posttranscriptional modifications in 16S and 23S rRNAs of the archaeal hyperthermophile *Sulfolobus sofataricus*. *Journal of Bacteriology* **180**, 2883-2888 (1998).

66. Bruenger, E. et al. 5S rRNA modification in the hyperthermophilic archaea *Sulfolobus sofataricus* and *Pyrodictium occultum*. *FASEB Journal* **7**, 196-200 (1993).

67. Zueva, V.S., Mankin, A.S., Bogdanov, A.A., Thurlow, D.L. & Zimmermann, R.A. Occurrence and location of 7-methylguanine residues in small-subunit ribosomal RNAs form eubacteria, archaebacteria and eukaryotes. *FEBS Letters* **188**, 233-238 (1985).

68. Ofengand, J. & Bakin, A. Mapping the nucleotide resolution of pseudouridine residues in large subunit ribosomal RNAs from representative eukaryotes, prokaryotes, archaebacteria, mitochondria and chloroplasts. *Journal of Molecular Biology* **266**, 246-268 (1997).

69. Kowalak, J.A. Identities and Phylogenetic Comparisons of Posttranscriptional Modifications in 16 S Ribosomal RNA from Haloferax volcanii. *Journal of Biological Chemistry* **275**, 24484-24489 (2000).

70. Kirpekar, F., Hansen, L., Rasmussen, A., Poehlsgaard, J. & Vester, B. The Archaeon has Few Modifications in the Central Parts of its 23S Ribosomal RNA. *Journal of Molecular Biology* **348**, 563-573 (2005).

71. O'Farrell, H.C., Scarsdale, J.N. & Rife, J.P. Crystal Structure of KsgA, a Universally Conserved rRNA Adenine Dimethyltransferase in Escherichia coli. *Journal of Molecular Biology* **339**, 337-353 (2004).

72. O'Farrell, H.C. Recognition of a complex substrate by the KsgA/Dim1 family of enzymes has been conserved throughout evolution. *RNA* **12**, 725-733 (2006).

73. Pulicherla, N. et al. Structural and Functional Divergence within the Dim1/KsgA Family of rRNA Methyltransferases. *Journal of Molecular Biology* **391**, 884-893 (2009).

74. Ohashi, Z., Maeda, M., McCloskey, J.A. & Nishimura, S. 3(-3-amino-3-carboxypropyl)-uridine: a novel modified nucleoside isolated from *Escherichia coli* phenylalanine transfer ribonucleic acid. *Biochemistry* **13**, 2620-2625 (1974).

75. Nishimura, S., Taya, Y., Kuchino, Y. & Ohashi, Z. Enzymatic synthesis of 3-(3-amino-3-carboxypropyl) uridine in *Esscherichia coli* phenylalanine transfer RNA: transfer of the 3-amino-3-carboxypropyl group from S-adenosylmethionine. *Biochemical and Biophysical Research Communications* **57**, 702-708 (1974).

76. Brochier-Armanet, C., Boussau, B., Gribaldo, S. & Forterre, P. Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nature Reviews Microbiology* **6**, 245-252 (2008).

77. Ferry, J.G. & Kastead, K.A. in Archaea. Molecular and Cellular Biology. (ed. R. Cavicchioli) 288-314 (American Society for Microbiology Press, Washington, DC; 2007).

78. Taylor, C.D. & Wolfe, R.S. Structure and methylation of coenzyme M. *Journal of Biological Chemistry* **249**, 4879-4885 (1974).

79. Hao, B. A New UAG-Encoded Residue in the Structure of a Methanogen Methyltransferase. *Science* **296**, 1462-1466 (2002).

80. Srinivasan, G. Pyrrolysine Encoded by UAG in Archaea: Charging of a UAG-Decoding Specialized tRNA. *Science* **296**, 1459-1462 (2002).

81. Blight, S.K. et al. Direct charging of tRNA-CUA with pyrrolysine *in vitro* and *in vivo*. *Nature* **431**, 333-335 (2004).

82. Mahapatra, A. et al. Characterization of a Methanosarcina acetivorans mutant unable to translate UAG as pyrrolysine. *Molecular Microbiology* **59**, 56-66 (2006).

83. Barns, S.M., Delwiche, C.F., Palmer, J.D. & Pace, N.R. Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *Proceedings of the National Academy of Sciences* **93**, 9188-9193 (1996).

84. Huber, H. et al. A new phylum of Archaea represented by a nonsized hyperthermophilic symbiont. *Nature* **417**, 63-67 (2002).

85. Schneider, K.L., Pollard, K.S., Baertsch, R., Pohl, A. & Lowe, T.M. The UCSC Archaeal Genome Browser. *Nucleic Acids Research* **34**, D407-D410 (2006).

86. Chan, P.P., Holmes, A.D., Smith, A.M., Tran, D. & Lowe, T.M. in Nucleic Acids Research, Vol. 402012).

87. Spang, A. et al. Distinct gene set in two different lineages of ammonia-oxidizing archaea supports the phylum Thaumarchaeota. *Trends in Microbiology* **18**, 331-340 (2010).

88. Paper, W. et al. Ignicoccus hospitalis sp. nov., the host of 'Nanoarchaeum equitans'. *International Journal of Systematic and Evolutionary Microbiology* **57**, 803-808 (2007).

89. Edwards, K.J., Bond, P.L., Gihring, T.M. & Banfield, J.F. An archaeal iron-oxidizing extreme acidophile important in acid mine drainage. *Science* **287**, 1796-1799 (2000).

90. Okibe, N., Gericke, M., Hallberg, K.B. & Johnson, D.B. Enumeration and Characterization of Acidophilic Microorganisms Isolated from a Pilot Plant Stirred-Tank Bioleaching Operation. *Applied and Environmental Microbiology* **69**, 1936-1943 (2003).

91. Macalady, J.L. et al. Tetraether-linked membrane monolayers in Ferroplasma spp: a key to survival in acid. *Extremophiles* **8**, 411-419 (2004).

92. Bond, P.L., Druschel, G.K. & Banfield, J.F. Comparison of acid mine drainage microbial communities in physically and geochemically distinct ecosystems. *Applied and Environmental Microbiology* **66**, 4962-4971 (2000).

93. Moses, C.O., Nordstrom, D.K., Herman, J.S. & Mills, A.L. Aqueous pyrite oxidation by dissolved oxygen and by ferric iron. *Geochimica et Cosmochimica Acta* **51**, 1561-1571 (1987).

94. Walsby, A.E. A Square Bacterium. *Nature (London)* **283**, 69-71 (1980).

95. Burns, D., Camakaris, H., Janssen, P. & Dyallsmith, M. Cultivation of Walsby's square haloarchaeon. *FEMS Microbiology Letters* **238**, 469-473 (2004).

96. Bolhuis, H., Poele, E.M.t. & Rodriguez-Valera, F. Isolation and cultivation of Walsby's square archaeon. *Environmental Microbiology* **6**, 1287-1291 (2004).

97. Elkins, J.G. et al. A korarchaeal genome reveals insights into the evolution of the Archaea. *Proceedings of the National Academy of Sciences* **105**, 8102-8107 (2008).

98. Auchtung, T.A., Takacs-Vesbach, C.D. & Cavanaugh, C.M. 16S rRNA Phylogenetic Investigation of the Candidate Division "Korarchaeota". *Applied and Environmental Microbiology* **72**, 5077-5082 (2006).

99. Reigstad, L.J., Jorgensen, S.L. & Schleper, C. Diversity and abundance of Korarchaeota in terrestrial hot springs of Iceland and Kamchatka. *The ISME Journal* **4**, 346-356 (2009).

100. Huber, H. et al. *Ignicoccus* gen. nov., a novel genus of hyperthermophilic chemolithoautotrophic *Archaea*, represented by two new species, *Ignicoccus islandicus* sp. nov. and *Ignicoccus pacificus* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* **50**, 2093-2100 (2000).

101. Waters, E. The genome of Nanoarchaeum equitans: Insights into early archaeal evolution and derived parasitism. *Proceedings of the National Academy of Sciences* **100**, 12984-12988 (2003).

102. Brochier, C., Gribaldo, S., Zivanovic, Y., Confalonieri, F. & Forterre, P. *Genome Biology* **6**, R42 (2005).

103. Jahn, U., Summons, R., Sturt, H., Grosjean, E. & Huber, H. Composition of the lipids of Nanoarchaeum equitans and their origin from its host Ignicoccus sp. strain KIN4/I. *Archives of Microbiology* **182**, 404-413 (2004).

104. Randau, L., Munch, R., Hohn, M.J., Jahn, D. & Soll, D. *Nanoarchaeum equitans* creates functional tRNAs from separate genes for their 5'- and 3'-halves. *Nature* **433**, 537-541 (2005).

105. Randau, L. The heteromeric Nanoarchaeum equitans splicing endonuclease cleaves noncanonical bulge-helix-bulge motifs of joined tRNA halves. *Proceedings of the National Academy of Sciences* **102**, 17934-17939 (2005).

106. Randau, L., Pearson, M. & Soll, D. The complete set of tRNA species in Nanoarchaeum equitans. *FEBS Letters* **579**, 2945-2947 (2005).

107. Huber, H., Hohn, M.J., Stetter, K.O. & Rachel, R. The phylum Nanoarchaeota: Present knowledge and future perspectives of a unique form of life. *Research in Microbiology* **154**, 165-171 (2003).

108. Könneke, M. et al. Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**, 543-546 (2005).

109. Walker, C.B. et al. *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. *Proceedings of the National Academy of Sciences* **107**, 8818-8823 (2010).

110. Hatzenpichler, R. et al. A moderately thermophilic ammonia-oxidizing crenarchaeote from a hot spring. *Proceedings of the National Academy of Sciences* **105**, 2134-2139 (2008).

111. Nunoura, T. et al. Insights into the evolution of Archaea and eukaryotic protein modifier systems revealed by the genome of a novel archaeal group. *Nucleic Acids Research* **39**, 3204-3223 (2011).

112. Ginzburg, M., Sachs, L. & Ginzburg, B.Z. Ion metabolism in *Halobacterium*. *Journal of General Microbiology* **55**, 187-207 (1970).

113. Volcani, B.E. in Papers collected to commemorate the 70th anniversary of Dr. Chaim Weizmann. 71-85 (Collective volume. Daneil Sieff Research Institute, Rehovoth; 1944).

114. Oren, A., Ginzburg, M., Ginzburg, B.Z., Hochstein, L.I. & Volcani, B.E. *Haloarcula marismortui* (Volcani) sp. nov., nom. rev., an extremely halophilic bacterium from the Dead Sea. *International Journal of Bacteriology* **40**, 209-210 (*1990*).

115. Pyatibratov, M.G. et al. Alternative flagellar filament types in the haloarchaeon *Haloarcula marismortui*. *Canadian Journal of Microbiology* **54**, 835-844 (2008).

116. Christian, J.H.B. & Waltho, J.A. Solute concentratoins within cells of halo-philic and non-halophilic bacteria. *Biochimica et Biophysica Acta* **65**, 506-508 (1962).

117. Shevack, A., Gewitz, H.S., Hennemann, B., Yonath, A. & Wittmann, H.G. Characterization and crystallization of ribosomal particles from *Halobacterium marismortui*. *FEBS Letters* **184**, 68-71 (1985).

118. Ban, N. et al. A 9A resolution X-ray crystallographic map of the large ribosomal subunit. *Cell* **93**, 1105-1115 (1998).

119. Zhao, P. The 2009 Nobel Prize in chemistry: Thomas A. Steitz and the structure of the ribosome. *Yale Journal of Biology and Medicine* **84**, 125-129 (2011).

120. Ban, N. et al. Placement of protein and RNA structures into a 5A-resolution map of the 50S ribosomal subunit. *Nature* **400**, 841-847 (1999).

121. Ban, N. The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4 A Resolution. *Science* **289**, 905-920 (2000).

122. Hildebrand, E. & Dencher, N. Two photosystems controlling behavioural responses of *Halobacterium halobium*. *Nature* **257**, 46-48 (1975).

123. Fu, H.Y. et al. A Novel Six-Rhodopsin System in a Single Archaeon. *Journal of Bacteriology* **192**, 5866-5873 (2010).

124. Mukohata, Y., Ihara, K., Tamura, T. & Sugiyama, Y. Halobacterial rhodopsins. *Journal of Biochemistry* **125**, 649-657 (1999).

125. Spudich, J.L. The multitalented microbial sensory rhodopsins. *Trends in Microbiology* **14**, 480-487 (2006).

126. Oesterhelt, D. & Stoeckenius, W. Rhodopsin-like protein from the purple membrane of *Halobacterium halobium*. *Nature New Biology* **233**, 149-152 (1971).

127. Matsuno-Yagi, A. & Mukohata, Y. ATP synthesis linked to light-dependent proton uptake in a red mutant strain of *Halobacterium* lacking baceriorhodopsin. *Archives of Biochemistry and Biophysics* **199**, 297-303 (1980).

128. Schobert, B. & Lanyi, J.K. Halorhodopsin is a light-driven chloride pump. *Journal of Biological Chemistry* **257**, 10306-10313 (1982).

129. Bogomolni, R.A. & Spudich, J.L. Identification of a third rhodopsin-like pigment in phototactic *Halobacterium halobium*. *Proceedings of the National Academy of Sciences* **79**, 6250-6254 (1982).

130. Spudich, J.L. & Bogomolni, R.A. Mechanism of colour discrimination by a bacterial sensory rhodopsin. *Nature* **312**, 509-513 (1984).

131. Takahashi, T., Mochizuki, Y., Kamo, N. & Kobatake, Y. Evidence that the long-lifetime photointermediate of s-rhodopsin is a receptor for negative phototaxis in halobacterium halobium. *Biochemical and Biophysical Research Communications* **127**, 99-105 (1985).

132. Bustin, S.A. et al. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clinical Chemistry* **55**, 611-622 (2009).

133. Robinson, J.L. et al. Growth Kinetics of Extremely Halophilic Archaea (Family Halobacteriaceae) as Revealed by Arrhenius Plots. *Journal of Bacteriology* **187**, 923-929 (2005).

134. Ginzburg, M. The unusual membrane permeability of two halophilic unicellular organisms. *Biochimica et Biophysica Acta* **173**, 370-376 (1969).

135. Tehei, M. et al. Neutron scattering reveals extremely slow cell water in a Dead Sea organism. *Proceedings of the National Academy of Sciences* **104**, 766-771 (2007).

136. Oren, A. Bioenergetic aspects of halophilism. *Mircobiology and Molecular Biology Reviews* **63**, 334-348 (1999).

137. Meury, J. & Kohiyama, M. ATP is required for K+ active transport in the archaebacterium *Haloferax volcanii*. *Archives of Microbiology* **151**, 530-536 (1989).

138. Rodríguez-Valera, F., Ruiz-Berraquero, F. & Ramos-Cormenzana, A. Isolation of extremely halophilic bacteria able to grow in defined inorganic media with single carbon sources. . *Journal of General Microbiology* **119**, 535-538 (1980).

139. Rodríguez-Valera, F., Juez, G. & Kushner, D.J. *Halobacterium mediterranei* spec. nov., a new carbohydrate-utilizing extreme halophile. *System Appl Microbiol* **4**, 369-381 (1983).

140. McLaren, C.E., Brittenham, G.M. & Hasselblad, V. Statistical and graphical evaluation of erythrocyte volume distributions. *American Journal of Physiology* **252**, H857-866 (1987).

141. Olesik, J.W. & Dziewatkoski, M.P. Time-resolved measurements of individual ion cloud signals to investigate space-charge effects in plasma mass spectrometry. *Journal of the American Society of Mass Spectrometry* **7**, 362-367 (1996).

142. Maher, W. et al. Measurement of trace elements and phosphorus in marine-animal and plant tissues by low-volume microwave digestion and ICP-MS. *Atomic Spectroscopy* **22**, 361-370 (2001).

143. Mulder, M.M., Teixeira de Mattos, M.J., Postma, P.W. & Dam, K. Energetic consequences of multiple potassium uptake systems in *Escherichia coli*. *Biochimica et Biophysica Acta* **851**, 223-228 (1986).

144. Bossemeyer, D. et al. Potassium transport protein TrkA of *Escherichia coli* is a peripheral membrane protein that requires other TrK gene products for attachment to the cytoplasmic membrane. *Journal of Biological Chemistry* **264**, 16403-16410 (1989).

145. Brown, G.R. & Cummings, S.P. Potassium uptake and retention by *Oceanomonas baumannii* at low water activity in the presence of phenol. *FEMS Microbiology Letters* **205**, 37-41 (2001).

146. Zhou, Y., Morais-Cabral, J.H., Kaufman, A. & Mackinnon, R. Chemistry of ion coordination and hydration revealed by a K+ channel-Fab complex at 2.0A resolution. *Nature* **414**, 43-48 (2001).

147. Higuchi, R., Dollinger, G., Walsh, P.S. & Griffith, R. Simultaneous amplificatoin and detection of specific DNA sequences. . *Biotechnology* **10**, 413-417 (1992).

148. Higuchi, R., Fockler, C., Dollinger, G. & Watson, R. Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Nature Biotechnology* **11**, 1026-1030 (1993).

149. Gibson, U.E., Heid, C.A. & Williams, P.M. A novel method for real time quantitative RT-PCR. *Genome Research* **6**, 995-1001 (1996).

150. Livak, K.J., Flood, S.J.A., Marmaro, J., Giusti, W. & Deetz, K. Oligonucleotides with fluorescent dyes at opposite end provide a quenched probe system useful for detecting PCR product and nucleic acid hybridizatoin. *Genome Research* **4**, 357-362 (1995).

151. Holland, P.M., Abramson, R.D., Watson, R. & Gelfand, D.H. Detection of specific polymerase chain reaction product by utilizing the 5' - 3' exonuclease activity of *Thermus aquaticus* DNA polymerase. *Proceedings of the National Academy of Sciences* **88**, 7276-7280 (1991).

152. Bustin, S.A. Absolute quantificatoin of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology* **25**, 169-193 (2000).

153. Gygi, S.P., Rochon, Y., Franza, B.R. & Aebersold, R. Correlation between protein and mRNA abundance in yeast. *Molecular and Cellular Biology* **19**, 1720-1730 (1999).

154. Bustin, S.A. & Nolan, T. Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *Journal of Biomolecular Techniques* **15**, 155-166 (2004).

155. Klug, G., Evguenieva-Hackenberg, E., Omer, A.D., Dennis, P.P. & Marchfelder, A. in Archaea. Molecular and Cellular Biology. (ed. R. Cavicchioli) 158-174 (American Society for Microbiology Press, Washington, DC; 2007).

156. Dai, L. ORF-less and reverse-transcriptase-encoding group II introns in archaebacteria, with a pattern of homing into related group II intron ORFs. *RNA* **9**, 14-19 (2003).

157. Macki, G.A. Ribonuclease E is a 5'-end-dependent endonuclease. *Nature* **395**, 720-723 (1998).

158. Kushner, S.R. mRNA Decay in Escherichia coli Comes of Age. *Journal of Bacteriology* **184**, 4658-4665 (2002).

159. Brown, J.W. & Reeve, J.N. Polyadenylated, noncapped RNA from the archaebacterium *Methanococcus vanielii*. *Journal of Bacteriology* **163**, 909-917 (1985).

160. Brown, J.W. & Reeve, J.N. Polyadenylated RNA isolated from the archaebacterium *Halobacterium halobium*. *Journal of Bacteriology* **166**, 686-688 (1986).

161. Leininger, S. et al. Archaea predominate among ammonia-oxidizing prokaryotes in soils. *Nature* **442**, 806-809 (2006).

162. López-López, A., Benlloch, S., Bonfá, M., Rodríguez-Valera, F. & Mira, A. Intragenomic 16S rDNA Divergence in Haloarcula marismortui Is an Adaptation to Different Temperatures. *Journal of Molecular Evolution* **65**, 687-696 (2007).

163. Rawls, K.S., Yacovone, S.K. & Maupin-Furlow, J.A. GlpR Represses Fructose and Glucose Metabolic Enzymes at the Level of Transcription in the Haloarchaeon Haloferax volcanii. *Journal of Bacteriology* **192**, 6251-6260 (2010).

164. Bidle, K.A. Differential expression of genes influenced by changing salinity using RNA arbitrarily primed PCR in the archaeal halophile Haloferax volcanii. *Extremophiles* **7**, 1-7 (2003).

165. Labrenz, M. et al. Relevance of a crenarchaeotal subcluster related to Candidatus Nitrosopumilus maritimus to ammonia oxidation in the suboxic zone of the central Baltic Sea. *The ISME Journal* **4**, 1496-1508 (2010).

166. Lipscomb, G.L. et al. Natural Competence in the Hyperthermophilic Archaeon Pyrococcus furiosus Facilitates Genetic Manipulation: Construction of Markerless Deletions of Genes Encoding the Two Cytoplasmic Hydrogenases. *Applied and Environmental Microbiology* **77**, 2232-2238 (2011).

167. Baliga, N.S. Genome sequence of Haloarcula marismortui: A halophilic archaeon from the Dead Sea. *Genome Research* **14**, 2221-2234 (2004).

168. Dennis, P.P., Ziesche, S. & Mylvaganam, S. Transcription analysis of two diparate rRNA operons in the halophilic archaeon *Haloarcula marismortui*. *Journal of Bacteriology* **180**, 4804-4813 (1998).

169. Portnoy, V. et al. RNA polyadenylation in Archaea: not observed in Haloferax while the exosome polynucleotidylates RNA in Sulfolobus. *EMBO reports* **6**, 1188-1193 (2005).

170. Kent, W.J. BLAT - the BLAST-like alignment tool. *Genome Research* **12**, 656-664 (2002).

171. Karolchik, D. et al. The UCSC Table Browser data retrieval tool. *Nucleic Acids Research* **32**, D493-D496 (2004).

172. Taylor, S., Wakem, M., Dijkman, G., Alsarraj, M. & Nguyen, M. A practical approach to RT-qPCR—Publishing data that conform to the MIQE guidelines. *Methods* **50**, S1-S5 (2010).

173. Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. & Vandesompele, J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biology* **8**, R19 (2007).

174. Vandesompele, J. et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* **3**, research 0034-0034.0011 (2002).

175. Shannon, P.T., Reiss, D.J., Bonneau, R. & Baliga, N.S. The Gaggle: An open-source software system for integrating bioinformatics software and data sources. *BMC Bioinformatics* **7** (2006).

176. Bare, J.C., Shannon, P.T., Schmid, A.K. & Baliga, N.S. The Firegoose: two-way integration of diverse data from different bioinformatic web resources with desktop applications. *BMC Bioinformatics* **8** (2007).

177. Bare, J.C., Kolde, T., Reiss, D.J., Tenenbaum, D. & Baliga, N.S. Integration and visualization of systems biology data in context of the genome. *BMC Bioinformatics* **11** (2010).

178. Mongodin, E.F. The genome of Salinibacter ruber: Convergence and gene exchange among hyperhalophilic bacteria and archaea. *Proceedings of the National Academy of Sciences* **102**, 18147-18152 (2005).

Appendix

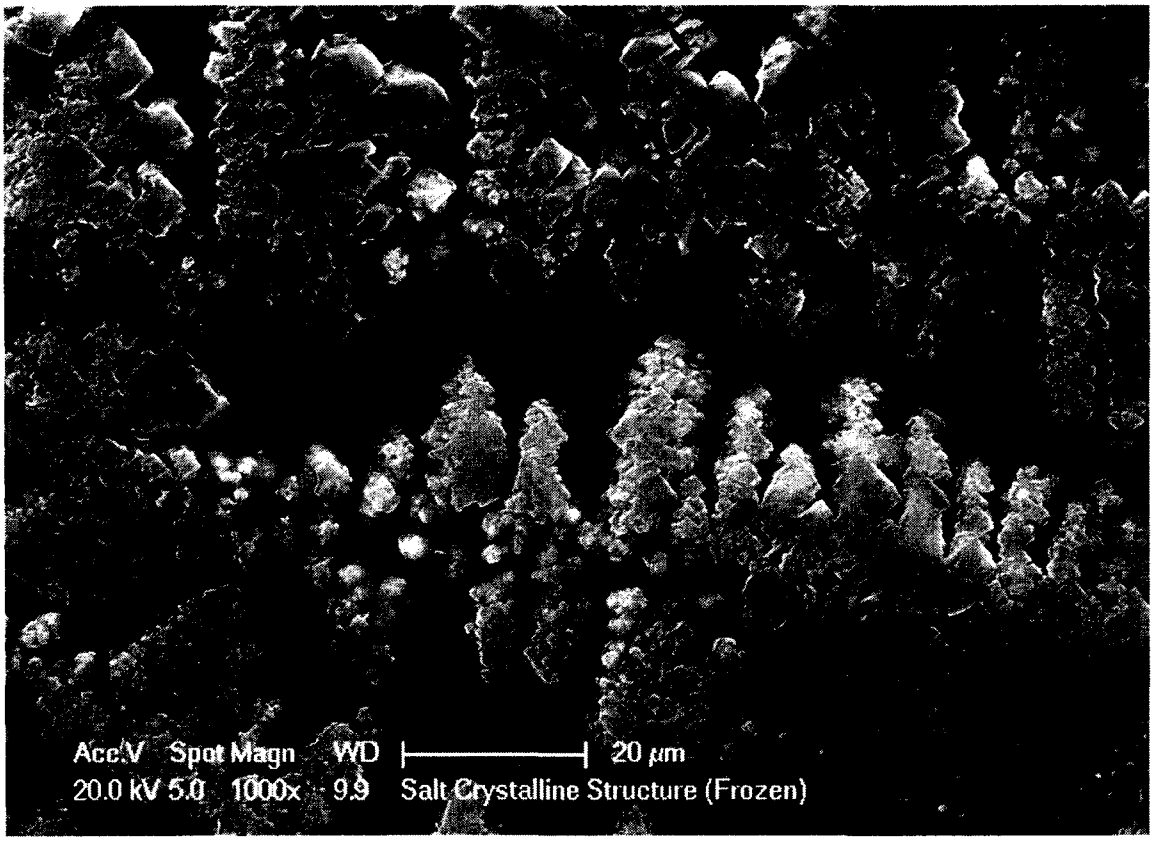

Figure A.1. Scanning electron micrograph of micro-crystalline salt structure. Crystals were formed during an attempt to image whole *Haloarcula marismortui* cells. Cells were pelleted by centrifugation and excess media was removed via pipette. Cell pellets were thinly spread across an imaging disk then submerged in liquid nitrogen for 30 seconds to solidify any remaining liquid material. The frozen ells were immediately gold plated under vacuum for 55 seconds at 45mA before imaging. Photograph was obtained using a Philips XL30 scanning electron microscope by Mr. Erwin Rehl, Department of Chemistry, University of Northern British Columbia.